

Advanced Second Language Segmental and Suprasegmental Acquisition

Kazuya Saito (Birkbeck College, University of London)

Abstract:

This chapter presents an overview on three fundamental questions in regard to advanced second language (L2) segmental and suprasegmental acquisition. They include (a) how L2 learners improve and attain advanced L2 speech abilities through extensive and intensive immersion in a target language speaking environment (i.e., *process*); (b) the extent to which their attained competence can approximate that of monolingual native speakers (i.e., *product*); and (c) what kinds of learners likely achieve such high-level L2 speech performance according to a range of learner-extrinsic factors (age of acquisition and testing, length of residence, quality and quantity of input) and learner-intrinsic factors (cognitive abilities, aptitude, motivation) (i.e., *individual differences*).

Key words:

Second language speech, pronunciation, segmentals, suprasegmentals, ultimate attainment, nativelikeness, advanced second language acquisition, individual differences, experience, age

Biography:

Kazuya Saito is a lecturer in the Department of Applied Linguistics and Communication at Birkbeck, University of London, UK. His research interests include how second language learners develop various dimensions of their speech in naturalistic settings; and how instruction can help optimize such learning processes in classroom contexts.

Please cite this paper as follows:

Saito, K. (in press). Advanced segmental and suprasegmental acquisition. In P. Malovrh & A. Benati (Eds.). *The handbook of advanced proficiency in second language acquisition*. Wiley Blackwell.

Introduction

Few second language acquisition (SLA) researchers would disagree with the fundamental idea that second language (L2) learners improve their linguistic performance as a function to increased experience, practice and interaction in the target language. For example, Flege (2016) pointed out that L2 learners' speech performance continues to develop, restructure and change over an extensive period of time (e.g., more than 10 years of length of residence [LOR]) before it becomes stable and plateaued, provided that their main language of communication is the L2 rather than their first language (L1). The ultimate attainment of these experienced L2 learners is typically referred to as "advanced" performance in the sense that it is significantly different from that of inexperienced L2 learners (e.g., LOR < 1 year) (McAllister, Flege, & Piske, 2002; Flege, Bohn, & Jang, 1997; Piske, McKay, & Flege, 2001). Among these experienced learners, it is believed that those with special individual difference profiles, such as an early age of acquisition (AOA) (Abrahamsson & Hyltenstam, 2009), high-level aptitude (Granena & Long, 2013) and motivation (Moyer, 2015), can reach near-nativelike proficiency. In the extensive literature on nativelikeness, which has investigated various dimensions of experienced and advanced L2 learners' linguistic performance, the acquisition of nativelike phonology seems to be particularly difficult compared to the acquisition of nativelike vocabulary and grammar (e.g., Granena & Long, 2013).

From a theoretical perspective, examining advanced L2 speech attainment—and distinguishing it from that of inexperienced L2 learners and native speakers—makes a substantial contribution to theory building in SLA. Unlike L1 acquisition, which generally results in nativelike proficiency across all linguistic domains of language (pronunciation, vocabulary and grammar), L2 learners' mental representations of language are built on their already-existing L1 systems (Best & Tyler, 2007; Flege, 2016). Thus, L2 systems are essentially different from those of monolinguals (Abrahamsson & Hyltenstam, 2009). Certain researchers (e.g., Cook, 2002) have claimed that any L2 phenomenon needs to be examined within non-native speakers themselves rather than via comparisons to a native speaker model. Focusing on beginner, intermediate and advanced L2 learners vs. native speakers, the growing number of empirical studies have examined the relatively difficult aspect of L2 speech learning—segmental and suprasegmental acquisition. Thus, it is high time that we summarize what comprises the process and product of advanced L2 speech learning and the extent to which L2 learners ultimately improve their segmental and suprasegmental proficiency relative to monolingual native speakers.

This chapter presents an overview of three fundamental questions regarding advanced L2 segmental and suprasegmental acquisition. These are: (a) how L2 learners improve and attain advanced L2 speech abilities through extensive and intensive immersion in a target language speaking environment (i.e., *process*); (b) the extent to which their attained competence can approximate that of monolingual native speakers (i.e., *product*); and (c) what kinds of learners likely achieve such high-level L2 speech performance according to a range of learner-extrinsic factors (age of acquisition and testing, length of residence, quality and quantity of input) and learner-intrinsic factors (cognitive abilities, aptitude, motivation) (i.e., *individual differences*). In this chapter, segmentals refer to individual consonantal and vocalic sounds in target language, whereas suprasegmentals are defined as "the structure that organizes sound" (Cutler, Dahan, & Donselaar, 1997, p. 142) subsuming a wide range of L2 phonetic phenomenon beyond segmentals, such as word stress, sentence stress, intonation, rhythm and fluency (including speed, breakdown and repair).

Process of Advanced-Level L2 Speech Attainment

In this subsection, we will first review previous literature investigating the developmental patterns found in the early and later stages of L2 speech learning (beginner → intermediate → advanced). Specifically, we will take a close look at how L2 learners develop new phonetic categories at the segmental and suprasegmental levels, and access them in perception and production, as they increase their amount of exposure to the target language.

Segmental Learning

With respect to the L1 acquisition literature, infants initially use computational strategies to detect prosodic patterns and then start recognizing words (Jusczyk, 1997). Importantly, after words are learned as whole phonological units, “the resulting increased vocabulary could result in sufficient pressure to fill in finer phonetic detail in the lexical representations in order to avoid confusion between similar sounding, known words” (Werker, & Tees, 1999, p. 523). This form of L1 phonetic development continues to take place up to adolescence (Walley & Flege, 1999). According to many L2 speech researchers, similar phenomena are observed in L2 phonetic development (e.g., Best & Tyler, 2007 for Perceptual Assimilation Model-L2; Bundgaard-Nielsen, Best, & Tyler, 2011 for Vocab Model; Flege, 1995 for Speech Learning Model; Walley, 2007 for Lexical Restructuring Model).

Once L2 learners become immersed in the target language community in naturalistic settings (e.g., study-abroad, immigration), they tend to prioritize the fast, efficient and robust recognition of, in particular, frequent and fundamental words (Nation & Webb, 2011); and emphasize producing them via intelligible and comprehensible pronunciation forms (Levis, 2005). Achieving basic L2 speaking proficiency is instrumental to successful social interaction with other native and non-native speakers in the early phases of L2 speech learning. According to corpus studies on L2 English, for example, these frequent and fundamental words comprise approximately 3,000 to 4,000 word families which cover 95% of vocabulary use in various daily communication settings (Adolphs & Schmitt, 2000); learning these can thus ensure a minimum and adequate understanding of various kinds of aural input (van Zeeland & Schmitt, 2012). As an attestation to this claim, Munro and Derwing (2008) longitudinally tracked the vowel production development of late 44 Mandarin and Slavic learners over their first year of residence in Canada. The results showed that much of their improvement was observed within the first six months of their immersion, especially for new vowel sounds within frequent words (e.g., [ɪ] in “bit” but not in “pit”).

As L2 learners increase their relevant experience and proficiency in the L2, however, they are forced to attend to fine-grained phonemic discrimination and identification (e.g., [i] vs. [ɪ], [p] vs. [b], [ɪ] vs. [l]). This is because these learners need to accurately comprehend and produce speech by drawing on a large lexicon containing many confusing minimal pairs (e.g., “beat” vs. “bit,” “pit” vs. “bit,” “read” vs. “lead”). At this stage, while L2 learners are sensitive primarily to word-sized units of L2 phonological information, they concurrently become more capable of detecting new sounds in the L2 input at a phonetic level. This phonetic-level restructuring is believed to lead L2 learners to create new phonetic categories and to generalize the newly-acquired phonetic knowledge from familiar to unfamiliar lexical contexts.

Turning our attention to recent L2 vocabulary research, there are many suggestions as to how much vocabulary is needed for L2 learners to trigger such phonetic-level restructuring (i.e., a transition from word to sound learning). When it comes to L2 English, for example, 6,000-7,000 word families are reported to account for 98% of lexical use in various discourse (e.g.,

Webb & Rodgers, 2011) and promote high-level comprehension of L2 aural texts (van Zeeland & Schmitt, 2012). Whereas knowing the most frequent 3,000-4,000 word families serves as a “minimum” requirement for beginner-to-intermediate L2 learners, knowing 6,000-7,000 word families could be considered as an “ideal” goal for advanced L2 learners (Schmitt, Cobb, Horst, & Schmitt, 2015).

Recently, Bundgaard-Nielsen et al. (2011) conducted a longitudinal research project with 31 Japanese learners of English with various vocabulary sizes (ranging between approximately 4,000-9,000 word families). Their longitudinal design is crucial, as the findings would directly relate to the causal relationship between vocabulary growth and the development of targetlike phonological behaviour. According to the results, the participants’ L2 English vowel identification patterns did not significantly change over a period of a few months. Rather, their vowel perception was strongly predicted by vocabulary size both at the beginning and end of the project. Notably, the participants with knowledge in the band of 6,000-7,000 word families demonstrated relatively superior L2 vowel perception. On the whole, the findings here in turn suggest that L2 learners’ increased segmental awareness likely takes place especially when their vocabulary size goes beyond the minimum threshold for daily conversations (3,000-4,000 word families) and reaches the lexical standard for advanced L2 learners (6,000-7,000 word families) (for the results of the production tests, see also Bundgaard-Nielsen, Best, Kroos, & Tyler, 2012).

Other L2 speech studies have also shown that more experienced L2 learners can attain more robust segmental competence as they become more capable of simultaneously processing word- and sound-sized units of L2 input. Owing to their relatively high awareness of individual consonants/vowels, these learners’ segmental performance may be less susceptible to the influence of lexical context. For example, Imai, Walley and Flege (2005) found that experienced Spanish learners’ word recognition in L2 English was consistent regardless of the different amount of lexical density (i.e., whether L2 words are phonologically similar or dissimilar to other words¹), although inexperienced Spanish learners’ perception abilities were significantly affected by lexical density across test items. Similarly, Flege, Frieda, Walley, and Randazza (1998) showed that highly experienced Spanish learners’ VOT production in L2 English was not affected by any of the crucial lexical factors that are hypothesized to determine the quality of L2 pronunciation performance (e.g., familiarity, abstractness) (see also Bradlow & Pisoni, 1999).

In sum, these studies equally indicate the presence of a strong vocabulary-phonology link—a phenomenon also attested in L1 acquisition—in the context of late L2 speech learning. As Bundgaard et al. (2011) pointed out, it seems reasonable to assume that “even late SLA is analogous to L1 acquisition in that the learning mechanisms employed in L1 acquisition may be available to the late L2 learner” (p. 459).

Suprasegmental Learning

Similar to L2 segmental learning, cross-sectional and longitudinal studies alike have suggested that L2 learners continue to enhance different aspects of suprasegmental production—prosody (word and sentence stress, intonation) and fluency (breakdown, speed, repair)—at different learning rates over an prolonged period of residence in a target-language-speaking environment. Using a longitudinal research design, Derwing, Munro, Thomson, and Rossiter (2009) examined late Chinese and Slavic learners’ suprasegmental (prosody, fluency) performance in L2 English over their first two years of immersion in Canada. Much

¹ Neighborhood density (i.e., phonological similarity) between words can be determined (a) substitution (e.g., “late” vs. “rate”), (b) deletion (e.g., “late” vs. “ate”), and (c) addition (e.g., “late” vs. “plate”).

improvement was observed for breakdown fluency (the number of filled and unfilled pauses), but only within the first year of residence in Canada. In contrast, the participants continued to enhance the prosodic (vowel durations) and speed fluency (articulation and speech rate) dimensions of their L2 speech over the duration of the project.

The “quick” and “immediate” improvement in L2 fluency suggested here is consistent with the findings of recent study-abroad research. For instance, Mora and Valls-Ferrer (2012) showed that Spanish college students demonstrated significant improvement in L2 English fluency as operationalized by breakdown measures (the number and duration of pauses) and speed measures (speech and articulation rate) over one academic semester in the U.S. For similar results with L2 learners of Spanish and French, see Towell, Hawkins, and Bazergui (1996) and Segalowitz and Freed (2004), respectively.

Conversely, improvement in L2 prosody—word and sentence stress, intonation—could be considered as “gradual” and “slow” in nature. Saito (2015a) conducted a cross-sectional research to examine the spontaneous speech of Japanese learners of English with various LOR profiles in Canada (1 month to 13 years). According to the results of linguistically trained raters’ subjective judgements, the prosodic (word stress, intonation) and temporal (speech rate) qualities of the participants’ speech was significantly predicted by their LOR. In the context of late Korean learners of English, Trofimovich and Baker’s (2006) cross-sectional study also demonstrated the significant predictive power of LOR for L2 suprasegmental learning; specifically, the participants’ duration ratio between stressed and unstressed syllables (crucial for the assignment of adequate word stress) was significantly associated with their LOR (1-10 years).

Taken together, the findings here indicate that L2 learners may quickly improve, in particular, the breakdown fluency of their L2 speech (the number of filled/unfilled pauses) in the early phases of L2 speech learning. This is arguably because the acquisition of these temporal features is assumed to be relevant to the surface-level processing of L2 speech production (i.e., monitoring). Yet, a great deal of L2 experience may be needed for L2 learners to enhance, refine and internalize other areas of L2 suprasegmental performance (i.e., word stress, intonation, speed fluency)—features which are thought to impact the automatization of phonetic encoding and articulation processes during L2 speech production.

Perception and Production.

As L2 learners develop segmental and suprasegmental representations as a function of increased experience and proficiency, they develop increasingly fluid access to these systems in order to comprehend and produce an L2 in various processing modes (perception, controlled and spontaneous production). Slightly different proposals have been developed to explain the inter-relationships between perception and production. For example, certain researchers have assumed that phonetic and phonological representations could be stored in the brain based on perception (how they hear temporal and spectral dimensions of new sounds) (e.g., Flege, 1995, 2009, 2016). Others have stated that these representations have a more articulatory basis (how they produce new sounds by using the tongue, lips and jaw) (Best & Tyler, 2007). Although there is some evidence that L2 learners’ perception abilities may precede their production abilities (perception-first) (e.g., Flege et al., 1997) or vice versa (e.g., Sheldon & Strange, 1982), these researchers have at least agreed with the fundamental idea that L2 learners simultaneously improve their perception and production abilities due to their interconnected nature (Bundgaard et al., 2012).

Whereas L2 learners' processing of new L2 sounds in the perception phase is relatively automatic (Flege, 1993), it remains open to debate to what degree their productive use of new L2 sounds is linked to their developing L2 phonology system. According to Major's (2008) Ontogeny Phylogeny Model, L2 learners initially tend to substitute their own L1 counterparts for new L2 sounds. With increasing awareness of new L2 sounds, however, interlanguage pronunciation performance begins to demonstrate some universal characteristics, regardless of L1 background. For instance, universally, L2 learners tend to make more pronunciation errors in free speech tasks than formal word reading tasks. Rau, Chang, and Tarone's (2009) demonstrated that Chinese learners of English mispronounced [θ] more frequently in a picture description task than in word and sentence reading tasks. Similarly, Wu and Saito (2014) also found that Cantonese learners of Mandarin Chinese produced more tone errors when their speech was elicited via word and sentence reading than via picture description. This task effect in performance was believed to arise from the increased demands on linguistic processing, due to the lack of substantial planning time for the picture description tasks, compared to the more controlled reading tasks.

Such task effects (controlled production > spontaneous production) can be explained by the cognitive psychology literature, which posits that L2 learners produce output via a gradual transition from effortful to automatic use of newly-acquired L2 knowledge (e.g., DeKeyser, 2007). Very importantly, this line of SLA research has also demonstrated that L2 learners can carefully produce language (and even speed up retrieval) without fully integrating linguistic knowledge into their mental representations (for a detailed discussion, see Segalowitz, 2003). In the context of L2 speech production, L2 learners can consciously activate relevant articulatory gestures in order to produce these sounds at a controlled-speech level (e.g., word and sentence reading), where they are given a sufficient amount of time to access their explicit knowledge.

The developmental patterns of the early and later phases of L2 speech learning are summarized in Figure 1.

Figure 1

Summary of L2 Speech Learning (Early Phase → Later Phase)

	Characteristics of early phase	→	Characteristics of ultimate attainment
Segmentals	Intelligible pronunciation of frequent and familiar words	→	More nativelike and refined pronunciation of segmentals regardless of lexical context
Suprasegmentals	Much improvement as measured via breakdown fluency measures (pause frequency)	→	Continuous improvement in prosody and speed fluency (articulation rate)
Processing	Perception may precede production or vice versa. Spontaneous production is lacking in the early stages.	→	Perception matches production. Spontaneous production ability balances out over time.

Product of Advanced-Level L2 Speech Attainment (1750 words)

Many experienced learners (e.g., LOR > 10 years) ultimately reach relatively high-level L2 performance compared to inexperienced learners (e.g., LOR < 1 year). Yet, it remains controversial the extent to which such high-level L2 competence can be close to native speaker ability. In this subsection, we will review previous empirical studies comparing linguistic abilities of three crucial groups: (a) experienced and advanced L2 learners, (b) inexperienced L2 learners, and (c) monolingual native speakers. As such, I will aim to provide a rough index of the potentials (vs. inexperienced L2 learners) and limitations (vs. native speaker baseline) of advanced L2 speech learning.

Segmental Attainment

In the early L2 speech literature, there is ample evidence that bilinguals' speech performance is essentially and fundamentally different from that of monolinguals (e.g., Sebastián-Gallés & Soto-Faraco, 1999). More recently, L2 researchers have extensively examined whether and to what degree experienced and advanced L2 learners can ultimately improve their speech performance to reach a "near-nativelike" level. Near-nativelikeness, defined as "levels of nonnativeness that are nonperceivable in normal, everyday language use" (Abrahamsson & Hyltenstam, 2009, p. 294), has been the focus of much empirical investigation. Notably, the concept of near-nativelikeness is different from nativelikeness, where non-native speakers have become not only perceptibly, but also linguistically equivalent with monolingual native speakers.

One well-known theoretical model especially related to the early phase of L2 speech learning is Flege's SLM, which specifically highlights the "ultimate" attainment and "near-nativelikeness" of L2 segmental acquisition (Flege, 1995, 2009, 2016). According to the SLM, when L2 learners identify a new L2 sound as perceptually distant from the closest L1 counterpart, they create a new phonetic category and strive to maintain its phonetic contrast relative to the L1 counterpart. Building on the SLM, Flege and his colleagues (e.g., McAllister et al., 2002) also proposed that such learning difficulty can be further predicted by the extent to which target acoustic (and articulatory) features are shared between the L1 and L2 phonetic inventories at a spectral (e.g., differences in formant frequencies) and temporal level (e.g., differences in segmental and formant transition length).

Following this line of thought, the amount of learning difficulty can be predicted in conjunction with cross-linguistic L1/L2 distance:

- The incidence of near-nativelikeness can be high for the adjustment of acoustic and articulatory phonetic features already present in both the L1 and L2.
- Few learners can achieve the near-nativelike proficiency required for establishing and acquiring complexly new features.

Since the fundamental concept of the SLM draws on the influence of the L1 on L2 development, it could also refer to the role of age of acquisition (the first intensive exposure to target language) in advanced-level SLA:

- Compared to late bilinguals whose L1 systems are fully developed and interfere with L2 speech learning to a great degree, early bilinguals, who are to start acquiring L2 while

developing their L1, may exhibit more advanced L2 attainment, as the process and product of early SLA may be less susceptible to L1 influence.

Munro, Flege and MacKay (1996) analyzed the English vowel production of 240 experienced Italian learners who widely varied in their AOA profiles in Canada (1-23 years). The results of listener judgements and acoustic analyses demonstrated that L2 English vowel production of many early learners (AOA < 7-8 years) was perceptibly indistinguishable from monolingual native speaker baseline. Such near-nativelike performance was not observed among late learners (AOA = 10-21 years).

A number of studies have delved into the role of L1 phonetic systems in the acquisition of L2 vowels by late learners (AOA > 16 years). Flege et al. (1997) explored the vowel perception and production of experienced (LOR = 7.3 years) and inexperienced (LOR = 0.7 years) L2 learners of English with three different L1 backgrounds (German, Korean, and Mandarin). Whereas the results noted that experienced L2 learners generally demonstrated better performance than inexperienced L2 learners, few achieved nativelike performance for any vowels. Importantly, the experienced learners' attainment was uniquely related to their L1 phonetic categories (e.g., German learners overused the durational cue to differentiate the English [ɛ]-[æ] distinction, as the temporal information is critical for the L1 counterparts (German [ɛ]) vs. [ɛ:]).

Similarly, Schmid, Gilbers, and Nota (2014) acoustically examined late experienced Dutch learners' vowel production of the phonemic contrast between [ɛ] ("pet") and [æ] ("trap"). The results showed that a very small portion of the participants attained nativelike vowel performance, suggesting that "late bilinguals, even at very advanced levels of proficiency, may have difficulties with vowel categories that are not instantiated by their L1" (p. 149).

To probe the detailed nature of L1 influence on late L2 vowel acquisition, McAlister et al. (2002) examined how experienced L2 learners of Swedish (LOR = 17-24 years) with three different backgrounds (English, Estonian, and Spanish) differentially perceived and produced vowel quantity. The results showed that, whereas Estonian learners' performance attained near-nativelike proficiency, English learners significantly differed from native Swedish controls and outperformed Spanish learners. McAlister et al. attributed the different ratio of ultimate attainment to the relative importance of vowel quantity in the participants' L1 phonetic systems (Estonian > English > Spanish).

With respect to consonants, much research has exclusively focused on the acquisition of the non-native contrast between English [ɹ] and [l] by Japanese learners, as it is hypothesized to be one of the most difficult instances in L2 speech learning (for a review, see Bradlow, 2008). Since Japanese has only one alveolar tap [ɾ], which acoustically lies somewhere between English [ɹ], [l] and [d] (Hattori & Iverson, 2009), Japanese learners of English tend to make much effort to create new phonetic categories for both English sounds, resulting in much learning difficulty in perception and production.

Early children likely demonstrate similar learning trajectories to those of monolingual children (phonetic development throughout childhood), and are able to attain near-nativelike proficiency (Idemaru & Holt, 2013). In contrast, few late learners (AOA > 16 years) attain such high-level L2 speech performance. In the previous literature, even experienced late Japanese learners (LOR > 10 years) perceived the English [ɹ]-[l] contrast with 80-90% accuracy (e.g., Ingvalson, McClelland, & Holt, 2011) and produced highly intelligible English [ɹ] and [l] pronunciation forms (e.g., Flege, Takagi, & Mann, 1995), although their perception and

production were significantly different from monolingual native speakers' baseline (Flege, Takagi, & Mann, 1996).

Saito (Saito, 2013; Saito & Brajot, 2013) acoustically analyzed English [ɹ] production in the context of 150+ Japanese learners with varied LOR in Canada (1 month to 40 years). It was found that most of the participants with more than 1 year of LOR experience had already demonstrated nativelike proficiency in terms of two articulatory parameters of English [ɹ]—tongue retraction (i.e., lower second formant) and segmental/formant transition lengthening (> 50ms), both of which are available in the Japanese phonetic system (e.g., [y] vs. [w] for the retraction parameter; [i] vs. [i:] for the lengthening parameter). Several Japanese learners—especially those with extensive LOR experience (> 6 years)—successfully acquired the new articulatory parameters: the labial, palatal and pharyngeal constrictions (i.e., lower third formant). However, their production fell short of native speaker baselines.

Another well-researched topic in the previous nativelikeness literature is the acquisition of voice onset time in word-initial stops (e.g., [p]-[b]). Unlike Japanese learners' English [ɹ] and [l] acquisition, it is intriguing that many L2 learners (including even late L2 learners) have been reported to attain nativelike proficiency in terms of this feature (e.g., Abrahamsson, 2012; Schmid et al, 2014; but see Chapter 17 for details of Voice Onset Time in Advanced SLA). This is arguably because L2 learners tend to show less difficulty in acquiring new temporal (e.g., adjusting voice onset time between L1 and L2) compared to spectral (e.g., enhancing sensitivity to third formants to differentiate English [ɹ]-[l]) information (see also Bohn & Flege, 1992).

However, it is also important to point out that L2 learners do not necessarily master all areas of L2 temporal features according to the availability of the relevant phonetic cues in L1 inventories. Using acoustic analysis, Baker (2011) closely examined two acoustic features of late Korean learners' "word-final" stop production in English—vowel and closure duration. The results showed that most of the participants mastered nativelike vowel duration—a phonetic cue that may be often used to differentiate short and long L1 Korean vowels. However, they did not reach nativelike-level closure duration—a new phonetic cue that is not accessed in producing L1 Korean sounds.

Suprasegmental Attainment

Though fewer in number, certain studies have examined the ultimate attainment and near-nativelikeness of L2 suprasegmental acquisition. Guion (2005) examined highly experienced Korean learners' (LOR > 10 years) knowledge and production of L2 English word stress, finding that even early learners' performance (AOA = 1-6 years) was significantly different from a monolingual native speaker baseline. Guion argued that the lack of nativelikeness can be ascribed to the crosslinguistic difference: Word-level prominence in L2 English does not exist in the L1 Korean phonetic system. In the context of a similar population (early English-Korean bilinguals), Trofimovich and Baker (2007) investigated a range of suprasegmental features (stress timing, intonation, speech rate, pause frequency and duration) in L2 English speech. The results showed (a) that early learners attained near-nativelikeness in prosody-based suprasegmentals (stress timing and intonation); and (b) that the fluency-based suprasegmental measure (speech rate) was significantly different from a monolingual native speaker baseline, regardless of the participants' AOA profiles.

As for fluency, Lahmann, Steinkrauss, and Schmid (2016) recently conducted a large-scale study with 102 experienced German learners of English (LOR > 20 years), investigating the acoustic characteristics of various dimensions of L2 fluency—mean number of unfilled

pauses (breakdown), mean syllable duration (speed) and mean number of self-corrections (repair). Because Lahmann et al. (2016) did not include a native baseline group, the study does not allow us to speculate the nativelikeness of the participants' advanced fluency performance. Saito (2015b) demonstrated that even highly experienced late Japanese-English bilinguals' (LOR > 10 years) perceived fluency was substantially different from a native speaker baseline.

The comprehensive review above suggests several broad, tentative conclusions about the ultimate attainment and near-nativeness of L2 segmental and suprasegmental acquisition. These conclusions are summarized in Table 1.

Individual Differences in Advanced-Level L2 Speech Attainment

Second language speech learning is subject to a great deal of individual variability, indicating that certain (but not all) L2 learners can make the most of their relevant L2 experience and attain advanced-level performance. This subsection will provide a systematic review on those learner variables thought to play a crucial role in determining the incidence of advanced- and near-native level L2 segmental and suprasegmental proficiency attainment. Such affecting factors include L2 learners' age (age of acquisition and testing), experience (quality and quantity of input), cognition (attentional and inhibition control, L2 learning aptitude), and conation (learner motivation and willingness to communicate).

Age

Flege, Munro, and MacKay (1995) explored the pronunciation attainment (measured via perceived accentedness) of 240 long-term Italian residents in Canada (LOR > 20 years). The results found that much variance (i.e., 59%) in their foreign accentedness could be explained by age of acquisition—the first intensive exposure to target language. Such findings have been replicated in many other L1/L2 contexts (e.g., Abrahamsson & Hyltenstam, 2009; Flege, Yeni-Komshian, & Liu, 1999). L2 researchers have put forward a number of accounts for explaining such strong age effects on L2 pronunciation attainment.

First, given that L1 and L2 acquisition are assumed to take place in the same phonetic space in the brain, an earlier AOA would seem to indicate a weaker influence of the L1 on L2 speech learning (Flege, 1995). Second, early bilinguals tend to have many opportunities to use the L2, especially in the schooling context, whereas certain late bilinguals can choose to interact with speakers of their L1 community rather than native speakers of the target language (Flege & MacKay, 2004). Third, L2 learners gradually lose their access to the implicit language acquisition device by which they pick up an L2 through mere exposure in an automatic and effortless fashion (Abrahamsson, 2012). When L2 learners are exposed to the L2 after puberty, their language learning process could be characterized as rather explicit, effortful and incomplete—a concept termed as the Critical Period Hypothesis. Fourth, many cognitive functions (which are assumed to be instrumental to successful language acquisition) begin to decrease after the late 20's, such as working memory, attentional and inhibitory control, and speech sound processing (Birdsong, 2014).

Table 1. *Summary of the Ultimate Attainment and Near-Nativeness of Experienced L2 Learners' Segmental and Suprasegmental Performance*

	Target features	Incidence of near-nativeness	Examples
Segmentals	Readjusting acoustic and articulatory features present in the L1	All learners attain high-level proficiency given ample input.	<ul style="list-style-type: none"> • Tongue retraction in the acquisition of English /ɪ/ by Japanese learners • Durational differences in various L2 segmental acquisition (e.g., VOT in the acquisition of English stops by Spanish learners) • Vowel duration for the acquisition of English word-final stops by Korean learners
	Acquiring acoustic and articulatory features absent in the L1	The acquisition of high-level proficiency requires early age of acquisition.	<ul style="list-style-type: none"> • Spectral difference in L2 vowel acquisition • Simultaneous labial, alveolar and pharyngeal constrictions in the acquisition of English /ɪ/ by L1 Japanese learners • Consonant closure for the acquisition of English word-final stops by Korean learners
Suprasegmentals	Prosody	The acquisition of high-level proficiency requires early age of acquisition.	<ul style="list-style-type: none"> • Word/sentence stress and intonation
	Breakdown and repair fluency	All learners attain high-level proficiency given ample input.	<ul style="list-style-type: none"> • Pause and reformulation frequency
	Speed fluency	Attaining high-level proficiency is extremely difficult regardless of age of acquisition.	<ul style="list-style-type: none"> • Articulation and speech rate

Interestingly, such AOA effects have not been reported when it comes to L2 fluency attainment. Saito (2015b) found that late experienced Japanese learners of English (LOR > 10 years) successfully attained high-level perceived fluency regardless of their age of acquisition. Lahmann et al. (2016) further illustrated that the breakdown (the number of unfilled pauses) and repair (the number of reformulations) dimensions of L2 fluency could be related to L2 learners' age of testing (when they took the tests) rather than age of acquisition (when they arrived in an L2 speaking country). The findings here suggest that humans gradually lose their "retrieval" abilities for their linguistic competence as a function of increased age—a similar phenomenon in L1 acquisition, where aging and dysfluency are strongly correlated with each other (e.g., Burke & Shafto, 2004, 2008). In other words, whereas the quality of phonological representations (segmental and suprasegmental accuracy) is greatly determined by the timing of arrival in an L2 speaking environment (i.e., age effects), L2 learners' speech production and monitoring abilities (reflecting on breakdown and repair fluency) correlate with the timing of speech data collection (i.e., aging effects).

Quality and Quantity of Input

According to Flege's (2009) oft-cited definition, input is termed as "L2 vocal utterances the learner has heard and comprehended, including his own, regardless of whether these utterances have been produced correctly by L2 native speakers, or incorrectly by other non-native speakers of the L2" (p. 175). It has been shown that L2 learners likely exhibit continuous improvement in L2 speech performance in relation to the amount of input via interaction with native and other non-native speakers. However, little learning occurs when L2 learners lack such willingness to communicate, regardless of how long they stay in an L2 speaking environment (Flege & MacKay, 2004; Derwing & Munro, 2013).

However, what remains unclear and controversial is the role of different "types" of input in L2 speech learning (Flege, 2009). For instance, what kind of input is most optimal for speech learning? To what extent does such optimal input facilitate SLA? To date, L2 speech researchers have made tremendous efforts to track not only the quantity, but also quality of L2 input received over years of immersion. However, these studies have exclusively focused on immigrants and measured their L2 experience in a retrospective way (e.g., self reports of frequency of L1 and L2 use). As a remedy, other L2 speech researchers have conducted intensive phonetic training studies, controlling the quality and quantity of L2 input and making precise descriptions of how different types of L2 input enhance the nature of phonetic and phonological representations.

High variability input: Logan and his colleagues (e.g., Logan, Lively, & Pisoni, 1991) conducted a series of studies to test the effects of High Variability Phonetic Training (HVPT; i.e., intensive exposure to natural L2 tokens produced by many talkers) on the acquisition of the /ɪ/-/i/ contrast by Japanese learners of English. The results showed that (a) learners who received HVPT (30 minutes × 15 sessions = 7.5 hours) could identify the English [ɪ]-[i] contrast at post-test sessions more successfully than at pre-test session; (b) the gains resulting from HVST were generalized to novel lexical contexts and new talkers; and (c) improvement was sustainable for six months without any additional training. Finally, the extended HVPT (30 minutes × 45 sessions = 22.5 hours) allowed learners to achieve perfect generalization (Yamada, 1995) and transfer the learning to the production domains (Bradlow, Pisoni, Akahane-Yamada, & Tohkura, 1997). Similar findings have been reported in the case of L2 vowel acquisition (Thomson, 2010) and L2 suprasegmental acquisition (e.g., Wang, Jongman, & Sereno, 2003).

Acoustically enhanced input: Given that exaggerated, acoustically-enhanced speech stimuli are hypothesized to be facilitative of discriminating non-native contrasts as shown in the L1 acquisition literature (Kuhl, 2000), other researchers have investigated the acoustic enhancement technique. In McCandliss, Fiez, Protopapas, Conway and McClelland's (2002) training study, Japanese learners first received synthesized tokens where the acoustic difference between English [ɪ] and English [I] was enhanced, and gradually moved on to listening to natural speech tokens. The results showed that acoustically enhanced input significantly helped the Japanese learners improve their English [ɪ]-[I] perception abilities.

Socially interactive input: Many L1/L2 researchers have examined the role of social interaction in L2 speech learning. In the bilingual infant literature, Kuhl and her colleagues have empirically shown that learners tend to acquire new foreign sounds through reciprocal interaction with actual persons compared to mere exposure to televised or audio-only input. This is because live situations arouse motivation and attention towards provided speech information via eye gazing and gestures to objects of reference (Conboy & Kuhl, 2011). Further, sharing the communicative intentions of others helps learners code word and phonetic information from ongoing speech (Kuhl, Tsao, & Liu, 2003). In the case of L2 speech learning, recent studies have provided some evidence that adult learners who receive input through meaning-oriented interaction with teachers tend to show quicker and more robust improvement with even a limited amount of training time (several hours), compared to those who practice target sounds via computer in a decontextualized manner (Lee & Lyster, 2015; Saito & Wu, 2014). Taken together, these studies (with both early and late bilinguals) suggest that social interaction might also play a key role in facilitating adult L2 speech learning, arguably because the presence of interlocutors provides learners with more immediate, engaging, and motivating conditions for learning compared to mere exposure to computerized treatment.

Cognition

A growing number of researchers have begun to investigate whether, to what degree and why L2 learners with special cognitive abilities (e.g., working memory, processing speed, inhibition control) and language learning aptitude (e.g., associative and rote memory, phonemic decoding) can achieve high-level speech perception and production abilities. For example, it has been shown that certain learners with strong working memories and processing speed can potentially optimize the L2 learning experience, as they have the capacity to store and access a great deal of received input efficiently and effectively (Darcy, Park, & Yang, 2015). Possessing inhibitory control also allows L2 learners to selectively focus on the target input while ignoring other information not relevant to the successful completion of tasks (Darcy, Mora, & Daidone, 2016). In terms of language learning aptitude, Granena and Long (2013) found that late Chinese-Spanish bilinguals' (AOA > 16 years) pronunciation performance (perceived foreign accent) was significantly related to their phonemic decoding test scores (i.e., connecting foreign sounds with corresponding symbols), whereas such an aptitude-proficiency correlation was not found for early bilinguals (AOA < 16 years). For similar findings in classroom settings, see Saito (2017).

Motivation

Finally, another well-researched variable in the context of advanced L2 speech attainment is learners' motivation. Previous research has revealed a surprising amount of variation in the motivation-attainment link, (for a review, see Piske et al., 2001), with some studies showing that

L2 learners with highly advanced oral proficiency are likely to have a great deal of professional motivation (e.g., “to teach an L2 as university-level academic jobs”: Moyer, 1999), instrumental motivation (e.g., “to get a job and/or respect at work”: Flege et al., 1995), integrative motivation (e.g., “to have as many native speaking friends as possible”: Flege et al., 1999) and strong concern for L2 pronunciation accuracy (e.g., “to pronounce English without any L2 accents”: Bongaerts, Summeren, Planken, & Schils, 1997). However, others have failed to find such significant predictive power of motivation for developing successful L2 pronunciation ability (e.g., Oyama, 1976; Purcell & Suter, 1980). The confounding finding signals that the relationship between motivation and L2 oral proficiency development is complex, and that research investigating this link needs to elaborate valid methods for quantifying motivation in various L1/L2 contexts (Piske et al., 2001).

Conclusion and Future Directions

Given the theoretical and practical relevance of the topic, this chapter set out to synthesize the extensive literature on advanced L2 segmental and suprasegmental acquisition. As suggested by many researchers (e.g., Flege, 2009), L2 learners continuously improve their segmental and suprasegmental competence and ultimately reach advanced level proficiency as long as they regularly use the target language as their main medium of communication over long periods of time (e.g., LOR > 10 years). With respect to segmentals, interlanguage development can be characterized as a transition from vocabulary learning (processing lexical units of L2 input as a whole) to sound learning (high sensitivity to segmental aspects of L2 speech). With respect to suprasegmentals (prosody, fluency), L2 learners’ speech may first become fluent (i.e., a reduction in the number of pauses at the early phase of L2 speech learning). Subsequently, L2 learners may slowly enhance the prosodic (word stress, intonation) and temporal (speech rate) qualities of their speech over time, as the automatization of phonetic encoding and articulation requires a great amount of L2 experience. The final quality of their asymptotic performance can be greatly determined in accordance with the L1/L2 cross-linguistic distance (the readjustment of existing cues vs. the acquisition of new cues), learner-extrinsic factors (age of acquisition and testing; quantity and quality of input) and learner-intrinsic factors (cognitive abilities, aptitude, motivation).

To close, I would like to point out a range of issues worthy of future investigation. First, most of the studies reviewed in this chapter have elicited L2 speech performance via controlled speech tasks (e.g., word and sentence reading tasks). This is arguably because these measures allow researchers to focus on target sounds while controlling for other confounding factors (e.g., variation in lexicogrammatical use). Under such highly controlled speech task conditions, however, L2 learners can carefully monitor correct language use, which does not necessarily reflect real-life situations, where language is used in various social settings in multiple tasks under time pressure. To tap into the present state of L2 learners’ segmental/suprasegmental representations and processing abilities, it is important for future studies to adopt not only controlled but also spontaneous and conversational speaking tasks (cf. Piske, Flege, MacKay, & Meador, 2011).

Another promising direction for future research relates to the role of individual differences in advanced L2 speech learning. Although researchers have begun to highlight the relationship between cognition, conation and ultimate attainment, the concept of individual differences has dramatically changed over the past 20 years. As reviewed above, for instance, the research findings have thus far supported the strong predictive validity of L2 learners’ cognitive

abilities for their successful L2 speech learning in the long run, especially when it comes to late L2 learners who likely learn their L2 in an intentional, explicit and analytical fashion. More research is needed to examine the importance of cognitive abilities and aptitude for implicit and incidental language learning in both the early and late phases. Several test formats have been adopted for this purpose, such as serial reaction time (e.g., Granena, 2013), semantic priming (e.g., Linck et al., 2013) and phonological sequences (e.g., Speciale, Ellis, & Bywater, 2004). Such future studies will in turn shed light on the cognitive correlates of successful L2 speech learning throughout one's lifetime.

Although the current review did not find a strong motivation effect on L2 speech learning and attainment, this conclusion needs to be interpreted with caution. Over the past 20 years, many researchers have proposed different theoretical models to capture the multifaceted, complex and dynamic nature of motivation (e.g., Dörnyei, Henry, & Muir, 2015 for the L2 Motivational Self System). Drawing on Norton's (2000) notion of imagined communities, for example, Yashima (2002) developed the idea of "international posture" to explain the specific motivation profiles of Japanese learners of English who are primarily driven by their desire to participate in an imagined international community. Saito, Dewaele and Hanzawa (in press) provided longitudinal evidence that L2 learners with context-specific motivation (learning English as long-term preparation for their an uncertain future career in an imaginary international community) can improve the comprehensibility and intelligibility of their pronunciation forms (but not necessarily their accentedness). These motivated learners tend to have a high willingness to communicate, and to seek and increase the number of opportunities to interact with native (and non-native) speakers in the target language community (Derwing & Munro, 2013). It is presumed that the increased amount of L2 experience will, as a result, facilitate SLA in the long run (Flege, 2009).

References

- Abrahamsson, N. (2012). Age of onset and nativelike L2 ultimate attainment of morphosyntactic and phonetic intuition. *Studies in Second Language Acquisition*, 34, 187–214.
- Abrahamsson, N. & Hyltenstam, K. (2009). Age of acquisition and nativelikeness in a second language – listener perception vs. linguistic scrutiny. *Language Learning*, 59, 249–306.
- Adolphs, S., & Schmitt, N. (2003). Lexical coverage of spoken discourse. *Applied Linguistics*, 24, 425–438.
- Baker, W. (2011). Effects of age and experience on the production of English word-final stops by Korean speakers. *Bilingualism: Language and Cognition*, 13, 263–278.
- Best, C., & Tyler, M. (2007). Nonnative and second-language speech perception. In O. Bohn, & M. Munro (Eds.), *Language experience in second language speech learning: In honour of James Emil Flege* (pp. 13–34). Amsterdam: John Benjamins.
- Best, C., McRoberts, G., & Goodell, E. (2001). Discrimination of non-native consonant contrasts varying in perceptual assimilation to the listener's native phonological system. *Journal of the Acoustical Society of America*, 109, 775–794.
- Birdsong, D. (2014). Dominance and age in bilingualism. *Applied Linguistics*. doi: amu031.
- Bohn, O. S., & Flege, J. E. (1992). The production of new and similar vowels by adult German learners of English. *Studies in Second Language Acquisition*, 14, 131-158.
- Bongaerts, T., Van Summeren, C., Planken, B., & Schils, E. (1997). Age and ultimate attainment in the pronunciation of a foreign language. *Studies in second language acquisition*, 19, 447-465.
- Bradlow, A. R. (2008). Training non-native language sound patterns. In J. Hansen & M. Zampini (Eds.), *Phonology and second language acquisition* (pp. 287–308). Philadelphia, PA: John Benjamins.
- Bradlow, A., & Pisoni, D. (1999). Recognition of spoken words by native and non-native listeners: Talker-, listener- and item-related factors. *Journal of the Acoustical Society of America*, 106, 2074–2085.
- Bradlow, A. R., Pisoni, D. B., Akahane-Yamada, R., & Tohkura, Y. I. (1997). Training Japanese listeners to identify English /ɪ/ and /I/. IV. Some effects of perceptual learning on speech production. *The Journal of the Acoustical Society of America*, 101, 2299-2310.
- Bundgaard-Nielsen, R., Best, C., Kroos, C., & Tyler, M. (2012). Second language learners' vocabulary expansion is associated with improved second language vowel intelligibility. *Applied Psycholinguistics*, 33, 643–664.
- Bundgaard-Nielsen, R., Best, C., & Tyler, M. (2011). Vocabulary size is associated with second-language vowel perception performance in adult learners. *Studies in Second Language Acquisition*, 33, 433–461.
- Burke, D. M., & Shafto, M. A. (2004). Aging and language production. *Current directions in psychological science*, 13(1), 21-24.
- Burke, D. M., & Shafto, M. A. (2008). Language and aging. *The handbook of aging and cognition*, 3, 373-443.
- Conboy, B. T., & Kuhl, P. K. (2011). Impact of second-language experience in infancy: brain measures of first-and second-language speech perception. *Developmental Science*, 14, 242-248.
- Cook, V. (ed.) (2002). *Portraits of the L2 user*. Clevedon: Multilingual Matters Ltd.
- Cutler, A, Dahan, D, & van Donselaar, W. (1997). Prosody and the comprehension of spoken language: A literature review, *Language and Speech*, 40, 141–201.

- Darcy, I., Mora, J., & Daidone, D. (2016). The role of inhibitory control in second language phonological processing. *Language Learning*. Advanced online publication. doi: 10.1111/lang.12161
- Darcy, I., Park, H., & Yang, C. L. (2015). Individual differences in L2 acquisition of English phonology: The relation between cognitive abilities and phonological processing. *Learning and Individual Differences*, 40, 63-72.
- DeKeyser, R. (Ed.). (2007). *Practice in a second language: Perspectives from applied linguistics and cognitive psychology*. Cambridge, UK: Cambridge University Press.
- Derwing, T. M., & Munro, M. J. (2013). The Development of L2 Oral Language Skills in Two L1 Groups: A 7-Year Study. *Language Learning*, 63, 163-185.
- Derwing, T. M., Munro, M. M., Thomson, R. I., & Rossiter, M. J. (2009). The relationship between L1 fluency and L2 fluency development. *Studies in Second Language Acquisition*, 31, 533–557.
- Dörnyei, Z., Henry, A., & Muir, C. (2015). *Motivational currents in language learning: Frameworks for focused interventions*. Routledge.
- Flege, J. (1993). Production and perception of a novel, second-language phonetic contrast. *Journal of the Acoustical Society of America*, 93, 1589–1608.
- Flege, J. (1995). Second language speech learning. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 233–277). Timonium, MD: York Press.
- Flege, J. E. (2009). Give input a chance! In T. Piske & M. Young-Scholten (Eds.), *Input matters in SLA* (pp.175–190). Clevedon: Multilingual Matters.
- Flege, J.E. (2016). *The role of phonetic category formation in second language speech acquisition*. Eight International Conference on Second Language Speech, 10 June, 2016, Aarhus University, Denmark.
- 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, 437–470.
- Flege, J., Frieda, E., Walley, A., & Randazza, L. (1998). Lexical factors and segmental accuracy in second-language speech production. *Studies in Second Language Acquisition*, 20, 155–188.
- Flege, J. E., & MacKay, I. R. (2004). Perceiving vowels in a second language. *Studies in second language acquisition*, 26, 1-34.
- Flege, J. E., Munro, M. J., & MacKay, I. R. (1995). Factors affecting strength of perceived foreign accent in a second language. *The Journal of the Acoustical Society of America*, 97, 3125-3134.
- Flege, J., Takagi, N., & Mann, V. (1995). Japanese adults learn to produce English /ɹ/ and /l/ accurately. *Language and Speech*, 38, 25–55.
- Flege, J., Takagi, N., & Mann, V. (1996). Lexical familiarity and English-language experience affect Japanese adults' perception of /ɹ/ and /l/. *Journal of Acoustical Society of America*, 99, 1161–1173.
- Flege, J. E., Yeni-Komshian, G. H., & Liu, S. (1999). Age constraints on second-language acquisition. *Journal of memory and language*, 41, 78-104.
- Granena, G. (2013). Individual differences in sequence learning ability and second language acquisition in early childhood and adulthood. *Language Learning*, 63, 665-703.

- Granena, G., & Long, M. H. (2013). Age of onset, length of residence, language aptitude, and ultimate L2 attainment in three linguistic domains. *Second Language Research*, 29, 311–343.
- Hattori, K., & Iverson, P. (2009). English /ɪ/-/I/ category assimilation by Japanese adults: Individual differences and the link to identification accuracy. *Journal of the Acoustical Society of America*, 125, 469–479.
- Idemaru, K., & Holt, L. L. (2013). The developmental trajectory of children's perception and production of English/ɪ/-/I/. *The Journal of the Acoustical Society of America*, 133, 4232–4246.
- Imai, S., Walley, A., & Flege, J. (2005). Lexical frequency and neighborhood density effects on the recognition of native and Spanish-accented words by native English and Spanish listeners. *Acoustical Society of America*, 117, 896–907.
- Ingvalson, E. M., Holt, L. L., & McCLELLAND, J. L. (2012). Can native Japanese listeners learn to differentiate/r-ɹ/ on the basis of F3 onset frequency? *Bilingualism: Language and Cognition*, 15, 255–274.
- Kuhl, P. K. (2000). A new view of language acquisition. *Proceedings of the National Academy of Sciences*, 97, 11850–11857.
- Kuhl, P. K., Tsao, F. M., & Liu, H. M. (2003). Foreign-language experience in infancy: Effects of short-term exposure and social interaction on phonetic learning. *Proceedings of the National Academy of Sciences*, 100, 9096–9101.
- Lahmann, C., Steinkrauss, R., & Schmid, M. S. (2016). Speed, breakdown, and repair: An investigation of fluency in long-term second-language speakers of English. *International Journal of Bilingualism*. doi: 1367006915613162.
- Lee, A. H., & Lyster, R. (2016). The effects of corrective feedback on instructed L2 speech perception. *Studies in Second Language Acquisition*, 38, 35–64.
- Levis, J. (2005). Changing contexts and shifting paradigms in pronunciation teaching. *TESOL Quarterly*, 39, 367–377.
- Logan, J. S., Lively, S. E., & Pisoni, D. B. (1991). Training Japanese listeners to identify English/ɪ/and/I: A first report. *The Journal of the Acoustical Society of America*, 89, 874–886.
- Major, R. (2008). Transfer in second language phonology: A review. In J. Hansen Edwards & M. Zampini (Eds.), *Phonology and Second Language Acquisition* (pp. 63–94). Amsterdam: John Benjamins.
- McAllister, R., Flege, J., & Piske, T. (2002). The influence of L1 on the acquisition of Swedish quantity by native speakers of Spanish, English and Estonian. *Journal of Phonetics*, 30, 229–258.
- McCandliss, B. D., Fiez, J. A., Protopapas, A., Conway, M., & McClelland, J. L. (2002). Success and failure in teaching the [r]-[l] contrast to Japanese adults: Tests of a Hebbian model of plasticity and stabilization in spoken language perception. *Cognitive, Affective, & Behavioral Neuroscience*, 2, 89–108.
- Mora, J. C., & Valls-Ferrer, M. (2012). Oral fluency, accuracy, and complexity in formal instruction and study abroad learning contexts. *TESOL Quarterly*, 46, 610–641.
- Moyer, A. (2015). Autonomy in second language phonology: Choice vs. limits. *Language Teaching*, 1–17.
- Munro, M. & Derwing, T. (2008). Segmental acquisition in adult ESL learners: A longitudinal study of vowel production. *Language Learning*, 58, 479–502.

- Munro, M., Flege, J., & MacKay, I. (1996). The effect of age of second-language learning on the production of English vowels. *Applied Psycholinguistics*, 17, 313–334.
- Nation, I.S.P., & Webb, S. (2011). *Researching and analyzing vocabulary*. Boston, MA: Heinle.
- Norton, B. (2000). *Identity and language learning: Gender, ethnicity and educational change*. Longman, Harlow, England.
- Oyama, S. (1976). A sensitive period for the acquisition of a nonnative phonological system. *Journal of Psycholinguistic Research*, 5, 261–283.
- Piske, T., Flege, J., MacKay, & Meador, D. (2011). Investigating native and non-native vowels produced in conversational speech. In M. Wrembel, M. Kul & K. Dziubalska-Kolaczyk (Eds.), *Achievements and perspectives in the acquisition of second language speech: New Sounds 2010* (pp. 195–205). Frankfurt am Main: Peter Lang.
- Piske, T., MacKay, I., & Flege, J. (2001). Factors affecting degree of foreign accents in an L2: A review. *Journal of Phonetics*, 29, 191–215.
- Purcell, E. T., & Suter, R. W. (1980). Predictors of pronunciation accuracy: A reexamination. *Language Learning*, 30, 271–287.
- Rau, D., Chang, A., & Tarone, E. (2009). Think or sink: Chinese learners' acquisition of the voiceless interdental fricative. *Language Learning*, 59, 581–621.
- Saito, K. (2013). Age effects on late bilingualism: The production development of /ɹ/ by high-proficiency Japanese learners of English. *Journal of Memory and Language*, 69, 546–562.
- Saito, K. (2015a). Experience effects on the development of late second language learners' oral proficiency. *Language Learning*, 65, 563–595.
- Saito, K. (2015b). The role of age of acquisition in late second language oral proficiency attainment. *Studies in Second Language Acquisition*, 37, 713–743.
- Saito, K. (2017). Effects of sound, vocabulary and grammar learning aptitude on adult second language oral proficiency in foreign language classrooms. *Language Learning*, 67, 665–693.
- Saito, K., & Brajot, F. (2013). Scrutinizing the role of length of residence and age of acquisition in the interlanguage pronunciation development of English /ɹ/ by late Japanese bilinguals. *Bilingualism: Language and Cognition*, 16, 847–863.
- Saito, K., Dewaele, J.-M., & Hanzawa, K. (in press). The role of motivation in late second language speech learning in foreign language classrooms: A longitudinal study. *Language and Speech*.
- Saito, K., & Wu, X. (2014). Communicative focus on form and L2 suprasegmental learning: Teaching Cantonese learners to perceive Mandarin tones. *Studies in Second Language Acquisition*, 36, 647–680.
- Schmid, M. S., Gilbers, S., & Nota, A. (2014). Ultimate attainment in late second language acquisition: Phonetic and grammatical challenges in advanced Dutch–English bilingualism. *Second Language Research*, 30, 129–157.
- Schmitt, N., Cobb, T., Horst, M., & Schmitt, D. (2015). How much vocabulary is needed to use English? Replication of van Zeeland & Schmitt (2012), Nation (2006) and Cobb (2007). *Language Teaching*, 1–15.
- Sebastián-Gallés, N., & Soto-Faraco, S. (1999). Online processing of native and non-native phonemic contrasts in early bilinguals. *Cognition*, 72, 111–123.
- Segalowitz, N. (2003) Automaticity and second languages. In C. Doughty & M. Long (Eds.), *The handbook of second language acquisition* (pp. 382–408). Oxford: Blackwell Publishers.

- Segalowitz, N., & Freed, B. F. (2004). Context, contact, and cognition in oral fluency acquisition: Learning Spanish in at home and study abroad contexts. *Studies in Second Language Acquisition*, 26, 173–199.
- Sheldon, A., & Strange, W. (1982). The acquisition of /ɪ/ and /I/ by Japanese learners of English: Evidence that speech production can precede speech perception. *Applied Psycholinguistics*, 3, 243–261.
- Speciale, G., Ellis, N. C., & Bywater, T. (2004). Phonological sequence learning and short-term store capacity determine second language vocabulary acquisition. *Applied Psycholinguistics*, 25, 293–321.
- Towell, R., Hawkins, R., & Bazergui, N. (1996). The development of fluency in advanced learners of French. *Applied Linguistics*, 17, 84–119.
- Trofimovich, P., & Baker, W. (2006). Learning second-language suprasegmentals: Effect of L2 experience on prosody and fluency characteristics of L2 speech. *Studies in Second Language Acquisition*, 28, 1–30.
- Trofimovich, P., & Baker, W. (2007). Learning prosody and fluency characteristics of L2 speech: The effect of experience on child learners' acquisition of five suprasegmentals. *Applied Psycholinguistics*, 28, 251–276.
- van Zeeland, H., & Schmitt, N. (2012). Lexical coverage in L1 and L2 listening comprehension: The same or different from reading comprehension? *Applied Linguistics*, 34, 457–479.
- Walley, A. (2007). Speech learning, lexical reorganization, and the development of word recognition by native and non-native English speakers. In O. Bohn, & M. Munro (Eds.), *Language experience in second language speech learning: In honour of James Emil Flege* (pp. 315–330). Amsterdam: John Benjamins.
- Walley, A. C., & Flege, J. E. (1999). Effect of lexical status on children's and adults' perception of native and non-native vowels. *Journal of Phonetics*, 27, 307–332.
- Wang, Y., Jongman, A., & Sereno, J. A. (2003). Acoustic and perceptual evaluation of Mandarin tone productions before and after perceptual training. *The Journal of the Acoustical Society of America*, 113, 1033–1043.
- Webb, S., & Rodgers, M. P. H. (2009). The lexical coverage of movies. *Applied Linguistics*, 30, 407–427.
- Werker, J. F., & Tees, R. C. (1999). Influences on infant speech processing: Toward a new synthesis. *Annual Review of Psychology*, 50, 509–535.
- Wu, X., & Saito, K. (2013). Difficulty in the acquisition of Mandarin high level and high falling tones by Cantonese learners. *The Journal of the Acoustical Society of America*, 134, 4106.
- Yamada, R. A. (1995). Age and acquisition of second language speech sounds: Perception of American English /p/ and /l/ by native speakers of Japanese. In W. Strange (ed.), *Speech perception and linguistic experience: Issues in cross-language research*, Timonium, MD: York Press, 305–320.
- Yashima, T. (2002). Willingness to communicate in a second language: The Japanese EFL context. *The Modern Language Journal*, 86, 54–66.