Second Language Speech Perception and the Brain

YUE WANG

The theoretical framework of second language (L2) speech perception and processing is based on the nature–nurture interplay in language learning. Theories have traditionally been dominated by the critical-period hypothesis (CPH), which suggests a maturation-based constraint in learning, resulting from the loss of plasticity of the brain after puberty. Current views posit that L2 speech learning is influenced by linguistic experience, as later learning may be constrained by initial exposure to a native language (L1). Nevertheless, speech learning can improve as learners gain experience in the L2, as the mature human brain is believed to remain malleable. While the nature–nurture issues cannot be framed in all-or-none dichotomies, the debate is still current as to the extent to which speech processing and learning are influenced by linguistic experience, or reflect lower-level encoding of physical properties, as well as the extent to which they involve an interconnected system within the realm of cognitive processing. Specifically, research in L2 speech processing and learning has addressed these issues from three perspectives: (a) whether the processing of L2 speech employs specialized neural substrates and how it is influenced by experience with an L1 speech system; (b) how learning-induced plasticity is instantiated in the brain; and (c) how L2 speech learning patterns extend across lower-level sensory physical and higher-level cognitive processes.

Differences in L1 and L2 Speech Processing

To address whether L2 speech engages specialized neural mechanisms, empirical research has investigated the nature of the differences between L1 and L2 speech processing. Initial research on the laterality effects of native and non-native speech compared the processing of Thai lexical tones by native Thai listeners and by English listeners whose L1 was nontonal (Van Lancker & Fromkin, 1973). The task was dichotic listening, in which two different tones were presented simultaneously, one to each ear. Since the right ear is primarily connected to the left hemisphere and the left ear to the right hemisphere through contralateral pathways, a right-ear advantage (indicating left-hemisphere dominance) was expected if tones were processed as linguistic stimuli. The results showed that only the native Thai listeners exhibited a significant left-hemisphere dominance. A follow-up study with musically trained English listeners revealed no hemispheric dominance for tone either, indicating that the left-hemisphere dominance found in the Thai listeners was due to their linguistic experience rather than greater familiarity with pitch contrasts. Consistently, research with Mandarin Chinese tone revealed that while Chinese listeners processed tone predominantly in the left hemisphere, Norwegian listeners exhibited bilateral processing, despite their experience with Norwegian tone (Wang, Behne, Jongman, & Sereno, 2004). Taken together, these studies suggest that left-hemisphere specialization for speech sounds occurs only when they are part of the listener’s phonemic repertoire.

The behavioral differences for native and non-native speech processing have been revealed in neural-based studies. One widely used approach is to examine the event-related
brain potentials (ERP), by comparing the differences in the peak amplitude and latency of brain waves associated with the detection of differences in the stimuli. A particularly important ERP component in auditory speech processing is called mismatch negativity (MMN), consisting of a peak amplitude difference in the waveforms at around 150–200 ms post-stimulus onset upon detection of stimulus differences. Thus, the existence and magnitude of MMN may reflect the ability to distinguish speech sound contrasts. ERP studies have revealed that native and non-native speech processing may involve different sensory acoustic and linguistic levels, with native processing more susceptible to higher-level linguistic influence. For example, in the processing of Hindi stop consonants, an MMN was seen in Hindi but not English listeners, whereas the two groups did not differ in the ERP responses reflecting lower-level processing of acoustic signals (Sharma & Dorman, 2000). Similar processes have been shown with the processing of phonemic pitch contrasts versus within-category acoustic changes in pitch, where only the native group exhibited enhanced MMNs for phonemic differences, but response patterns to acoustic changes were the same for both native and non-native groups (Chandrasekaran, Krishnan, & Gandour, 2009b). Indeed, research has consistently revealed larger MMNs in response to the native speech contrasts than to those that are nonexistent in listeners’ native phonetic inventories. These findings indicate that listeners are more sensitive to phonetic features in an L1 phonetic context, demonstrating experience-dependent neural effects in native and non-native speech processing.

Additionally, neuroimaging approaches (such as functional magnetic resonance imaging, fMRI) have been used to examine how the difference between native and non-native speech processing is localized in the brain. These techniques measure regional changes in blood flow in response to changes in neuronal activity, thus enabling localization of particular mental processes. Research has shown that the differences between native and non-native speech processing involve a network of brain regions, which go beyond the classic language areas in the left hemisphere such as Broca’s and Wernicke’s areas. For example, for lexical tone and vowel duration, native listeners revealed left-hemisphere activation in the prefrontal, frontal (Broca’s), temporal (Wernicke’s), and parietal areas, while non-native listeners exhibited predominant right-hemisphere processing but a lesser degree of left-hemisphere involvement, even for those whose L1 was tonal. In contrast, for nonlinguistic pitch or duration patterns the native and non-native listeners exhibited similar patterns of bilateral processing in the frontal and parietal areas (Gandour et al., 2002). These findings agree with the behavioral and ERP results discussed above, indicating a language-specific top-down processing of pitch where non-native listeners’ experience with acoustic cues may not be generated to higher-order linguistic processing. Moreover, the processing of different L2 linguistic speech features may involve different neural correlates depending on the nature of corresponding features in the L1. For example, while native tone-language users processed tone predominantly in the left-hemisphere temporal cortex, they processed intonation in the homologous right-hemisphere regions, presumably due to the greater linguistic functional use of tone than of intonation. In contrast, non-native listeners did not reveal any differences for the two tasks since tone was not used in their L1s (Fournier, Gussenhoven, Jensen, & Hagoort, 2010). Likewise, when processing prosodic focus and intonation, Chinese listeners showed different cortical patterns for L1 (Mandarin) and L2 (English) focus but not intonation, presumably because the phonetic manifestations of focus rather than intonation differ in the L1 and L2 (Gandour et al., 2007).

Thus, the behavioral and brain data consistently reveal that, compared to L1, L2 speech processing may involve delayed latency and weaker sensitivity, as well as different and additional cortical regions, with the nature of these differences being determined by L2 users’ linguistic experience.
Experience- and Learning-Induced Plasticity

Given the differences between native and non-native speech processing, one subsequent question is the extent to which the human brain has the capacity to change with continuous L2 learning. Empirical research has shown that language-related cortical responses differ as a function of experience with an L2, depending on such factors as age of L2 acquisition (AOA), L2 proficiency, and exposure to and training in the L2.

First, age has been shown to affect hemispheric lateralization for L2 speech: Learners with an early AOA tend to involve more left-hemisphere processing than do late learners (e.g., Sussman, Franklin, & Simon, 1982). Moreover, ERP research has revealed changes in MMN patterns as a function of L2 proficiency, showing that advanced rather than naive Hungarian learners of Finnish approximated native Finnish MMN patterns for processing the Finnish vowel contrasts (Winkler et al., 1999). These findings indicate experience-induced cortical differences in L2 speech processing.

One widely used method to experimentally assess the plasticity of the brain in L2 speech learning is laboratory-based perceptual training. Initial behavioral studies have shown that, after about two weeks’ training, adult L2 learners can significantly improve their perception of L2 speech contrasts, suggesting that the adult human perceptual system still has the capacity to change. Consistently, ERP results have demonstrated that training can enhance the MMN responses (particularly in the left hemisphere) to L2 speech contrasts (Zhang et al., 2009). Moreover, changes in neural sensitivity due to training may be differentially affected by listeners’ L1. For example, after training with Thai tones, Chinese listeners (with a tonal L1), compared to English listeners (with a nontonal L1), exhibited a larger negativity at a delayed latency, indicating that tonal L1 users were more sensitive to linguistically relevant pitch differences (Kaan, Wayland, Bao, & Barkley, 2007). Training effects have also been exhibited in the spatial domain, showing that the cortical effects of learning an L2 involve both the expansion of preexisting language-related areas and the recruitment of additional cortical regions. After training with English /r–l/, Japanese learners’ improvements were associated with extended bilateral cortical and subcortical regions involved with auditory-articulatory mappings related to speech processing (Callan et al., 2003). Likewise, for American learners of Mandarin tones, the cortical instantiation of training effects involved increased and expanded activation in Wernicke’s and adjacent areas as well as the right-hemisphere regions homologous to Broca’s area (Wang, Sereno, Jongman, & Hirsch, 2003). Moreover, training may result in a higher degree of neural efficiency, focusing on the areas involved in native processing (Golestani & Zatorre, 2004). Particularly, more successful learners have shown increased activation in the left-hemisphere regions associated with L1 linguistic processing, indicating an association between the degree of neural plasticity and language learning (Wong, Perrachione, & Parrish, 2007).

In sum, these patterns indicate progressive cortical changes with training and increased proficiency, suggesting that cortical representations may be continuously shaped with experience and learning.

Across Acoustic, Linguistic, and Cognitive Domains

While L2 speech processing apparently involves language-specific brain mechanisms, it is also a part of an interconnected neural system entailing integrated mechanisms across sensory-motor, linguistic, and general cognitive domains.

As discussed earlier, L2 speech processing may to a large extent overlap with the processing of acoustic information. For example, while native processing of fricative contrasts consistently elicited an MMN in the left hemisphere, non-native processing involved
a shift in MMN amplitude from the right hemisphere for the initial frication noise to the left hemisphere for the later fricative–vowel transition, reflecting non-native listeners’ sensitivity to the acoustic differences of the stimuli (Lipski & Mathiak, 2007). Moreover, the weaker and later MMN responses in processing some L2 speech contrasts may have stemmed from the reduction of acoustic information available to L2 learners or to the difficult acoustic context that the neighboring sounds create (Frenck-Mestre, Meunier, Espesser, Daffner, & Holcomb, 2005). Additionally, listeners may seek alternative routes to process L2 speech by encoding the cortical regions used for articulatory-orosensory, primary auditory, and acoustic processing (Callan, Jones, Callan, & Akahane-Yamada, 2004).

As speech processing occurs concurrently with that of other linguistic domains, the processing of L2 phonetic information may uniquely interact with lexical, semantic, and prosodic processing. For example, while Spanish learners showed increased activation in the right-hemisphere regions for processing phonological information compared to semantic information in their L2 (English), the processing patterns for these two types of stimuli in their L1 did not differ, suggesting shared neural networks for L1 and L2 semantic processing but L2-specific phonological neural circuits (Pillai et al., 2003). Similarly, L1 and L2 phonological representations have been shown to differ in the mental lexicon: Spanish-dominant Spanish–Catalan bilinguals exhibited a larger phonological error-related ERP component in detecting words involving Catalan vowels, whereas they showed a larger semantic-related ERP component for words involving Spanish vowels (Sebastian-Galles, Rodriguez-Fornells, de Diego-Balaguer, & Diaz, 2006).

L2 speech processing may also involve cognitive processes such as memory, attention, and cognitive control. Greater attention and cognitive control have been claimed to be associated with additional activities in the prefrontal and anterior cingulated cortex. Likewise, better L2 attainment may be correlated with more optimal engagement of phonological working memory, indexed by greater activation in the left insula (Chee, Soon, Lee, & Pallier, 2004). Additionally, speech shares some neural mechanisms with music, given that pitch and temporal perception are fundamental to both. It is thus conceivable that music experience may facilitate speech learning. Indeed, for linguistic pitch (tone) processing in Mandarin, English musicians approximated the native Chinese patterns with larger MMNs than English non-musicians (Chandrasekaran, Krishnan, & Gandour, 2009a). Music experience has also been shown to facilitate the categorization of prosodic endings in an L2, as indexed by the earlier ERP responses for non-native learners with musical experience compared to non-musicians (Marques, Moreno, Castro, & Besson 2007). These findings imply that long-term experience in pitch across domains may result in cortical plasticity to pitch.

**Conclusion and Future Directions**

Taken together, empirical evidence cumulatively indicates that L2 speech processing involves experience-dependent but domain-general processes. On the one hand, the processing of L2 speech does encapsulate unique neural substrates under the influence of the L1 system. On the other hand, it does not just engage a single process or a single domain, but rather functions in the context of language processing in general, reflecting the integration of cortical networks across sensory-motor and cognitive levels. Moreover, L2 speech processing involves a dynamic process as a function of experience and learning. A convergence of data speaks against the maturation-based CPH, revealing neural plasticity involving changes in the temporal dynamics and cortical organizations of L2 representations in the brain.

Nevertheless, the nature of the CPH in L2 speech learning is still a matter of ongoing debate, as the brain data have only recently been systematically quantified with the advent
of neuroimaging techniques. It remains to be elucidated how cortical representations can be continuously shaped with learning. Future longitudinal studies tracing L2 speech learning trajectories are needed in order to define the agents of neural plasticity at various stages. Another direction for future research is to take into account individual differences. Although not yet extensively investigated, existing evidence indicates that the nature of individual processing patterns may trace the origin of speech processing as well as predicting learning success. All in all, L2 speech research addressing different stages of learning, different domains involved, and different factors affecting learning will lead us toward a complete picture in unraveling the neural mechanisms underlying L2 speech learning.

SEE ALSO: Brain Activity During Second Language Processing (ERP); Critical Period; Research Techniques and the Bilingual Brain; Second Language Representation in the Brain; Speech Perception

References


### Suggested Readings


