Second Language Acquisition

P Indefrey and D J Davidson, Max Planck Institute for Psycholinguistics and F. C. Donders Centre for Cognitive Neuroimaging, Nijmegen, The Netherlands

Introduction: Second Language Acquisition and the Brain

Speaking a second language (L2) means having acquired a new inventory of phonemes (speech sounds), new words, and new inflectional and syntactic rules. Because this new knowledge must be stored somewhere in the brain, it does not require neuroimaging experiments to conclude that a bilingual speaker's brain must be in some way different from that of a monolingual speaker. The question is rather: Can the difference between the brain of a bilingual speaker and a monolingual speaker tell us something about the functional organization of a bilingual language system? Quite different answers are possible, ranging from 'more of the same,' meaning that at all processing levels the neural substrates for first language (L1) and L2 processing completely overlap, to 'radically different,' as suggested by early hypotheses that an L2 relies on the right rather than the left hemisphere of the brain. In between, there are hypotheses suggesting that L2s rely predominantly on a subset of L1 processes, such as storage and declarative processing, but not rule or procedural processing. Such hypotheses predict that a subset of the brain regions involved in L1 processing, namely those subserving rule or procedural processing, should not be recruited for L2 processing.

Neuroimaging studies on bilingual speakers are thus potentially of high relevance for our understanding of the functional architecture of a bilingual language system. At the same time, they face considerable methodological problems. In experiments involving a task, differences in neural activity patterns may be explained by general differences in task performance rather than the components of the task that the experiment was designed to assess. Given that most speakers are less proficient in their L2 than in their L1, however, equal task performance is rarely achieved. A second problem is that experimental variables such as 'age of L2 acquisition' are almost inevitably confounded with other variables such as 'L2 proficiency' or vice versa. Often, it becomes a matter of interpretation which of the observed neuronal activation differences are due to experimental variables and which are due to possible confounds. In this situation, comparisons across studies can be helpful because they may indicate which activation patterns or electrophysiological signatures are reliably found in certain tasks when performed in the L2, irrespective of particular L2 speaker characteristics. They may also show brain activity that is reliably found for subsets of studies testing particular types of L2 speakers.

This article therefore focuses on experimental tasks that have been used in a number of hemodynamic (i.e., positron emission tomography (PET) and functional magnetic resonance imaging (fMRI)) and electrophysiological studies of L2 processing, such that an assessment of the reliability of findings is possible. Where possible, we attempt to identify L2 speaker characteristics that seem to influence observed brain activation patterns.

Hemodynamic Correlates of L2 Acquisition

Hemodynamic studies on the neural correlates of L2 processing have been conducted on all levels of linguistic structure from phonology to discourse. At higher linguistic levels, most studies have investigated the processing of an L2 after years of learning rather than L2 acquisition as such. These studies provide evidence of changes in functional neuroanatomy that must have occurred at some time point during L2 acquisition rather than during the study. Experiments with true learning (longitudinal) designs are relatively rare and mostly found at the phonemic level.

Phonemic Learning

A number of studies have investigated the neural consequences of learning a novel phonemic contrast using within-subject training designs. These studies are very instructive because they illustrate a methodological problem of neurocognitive research on bilingualism and possible ways to overcome the problem. To detect neural correlates of phonemic learning, participants performing phoneme classification or recognition tasks were scanned with fMRI before and after a training phase (several hours spread over 1–4 weeks), during which they learned nonnative phonemic contrasts (e.g., Japanese speakers learning the English *r/l* contrast). All studies tested for differences between pre- and posttraining fMRI measurements. The number and anatomical location of regional changes in activation varied considerably across studies, but there were five regions that were

replicated across studies: the left and right posterior inferior frontal gyri, the left mid superior temporal gyrus (auditory cortex and planum temporale), the left superior parietal lobule, and the left caudate nucleus. According to the logic of the design, it is tempting to conclude that these regions are involved in the processing of the novel phonemic contrast. Note, however, that by definition successful training resulted in improved performance such that the two fMRI measurements were no longer comparable in this respect. Therefore, it cannot be excluded that the observed regions simply reflect unspecific effects, such as reduced attention or effort, lower error rates, and so forth. To be able to assess which of the observed activation differences were due to phonemic processing and which were unspecific effects of higher performance or even simple repetition effects, some studies included a control condition by training a second, easy phonemic contrast that was similar to a native language contrast. Because performance in this condition is sometimes even higher than that on the critical nonnative contrast, it is possible to interpret activation changes that only occurred in the critical condition as being specific for learning a nonnative contrast. When taking into account these additional data, commonalities between the findings of the different studies are further reduced.

Remaining regions that might be specifically recruited for learning a nonnative phonemic contrast are the left inferior frontal gyrus and the left caudate nucleus, but other data suggest that posttraining learner brain activation may be very much like native speaker brain activation. One possible reason for such seemingly discrepant findings is a potentially nonlinear relationship between learning and hemodynamic activation. The activation level of regions that are recruited for a difficult task (e.g., recognizing a nonnative contrast) may be initially high but fall off again after some level of performance is reached. This issue is further discussed later.

Word-Level Processing

The two most frequently used tasks for the study of word production are picture naming and word generation (typically generation of words that start with a given letter or fit semantically to a given stimulus word). Both tasks involve the core components of word production (lemma retrieval, word form retrieval, syllabification, and phonetic encoding) but differ with respect to the processes that are employed to come up with a lexical concept to be produced. No study comparing picture naming or word generation in the L1 and L2 of bilingual speakers has reported stronger activation when the task was performed in the native language. Stronger activation for the L2

has been reported in a small but reliable number of studies in the left (for word generation also the right) posterior inferior frontal lobe (**Figure 1**). Studies reporting stronger activation of this region in L2 picture naming had participants with later L2 onset and lower L2 proficiency compared to studies reporting no differences between L2 and L1 picture naming. Stronger activation of the left posterior inferior frontal gyrus was also found in a verb-generation study with quadri-lingual participants for languages that were learned later and spoken less proficiently. These results show that at a group level, L1 and L2 word production seem to engage the same cortical areas, and that L2 speakers with late onset or lower proficiency may recruit one or more of the areas more strongly. Note, however, that hemodynamic studies show all brain areas that are activated. They do not provide evidence of which brain areas are actually necessary for word production in individual speakers. Evidence for necessary areas comes from a study in which different cortical areas in fluent bilingual epilepsy patients were electrically stimulated prior to surgery. The patients performed an object-naming task in their first and their L2, and it was recorded whether electrical stimulation interfered with object naming or not. For every individual, this procedure resulted in a map of sites that were necessary for L1 picture naming, L2 picture naming, or both. Shared sites were found in all left peri-Sylvian regions. L1-specific sites were also found in both posterior frontal and temporal regions but more so in the frontal cortex. L2-specific sites were exclusively found in mid- to posterior temporal cortex and adjacent inferior parietal sites. These data show that although there may not be any L1- or L2-specific regions at a group level, there seem to be cortical sites in many individuals that are only necessary for word production in one of the languages. Moreover, the left inferior frontal cortex, which in hemodynamic studies has been found to be more strongly activated for L2 word production, does not contain L2-specific sites but, rather, L1-specific sites possibly subserving L1-specific processes. A possible interpretation might be that L2 speakers attempt to make use of processes that are in some way tailored to L1 word production (to the extent that L1, but not L2, word production becomes impossible without them). A good candidate for such a process seems to be postlexical syllabification, which is subject to language-specific constraints and engages Broca's area in L1 word production.

Most studies on bilingual word-level comprehension have used written stimuli and asked their subjects to perform some kind of semantic decision. Typically, control conditions involved nonlinguistic stimuli so that the observed activations may reflect
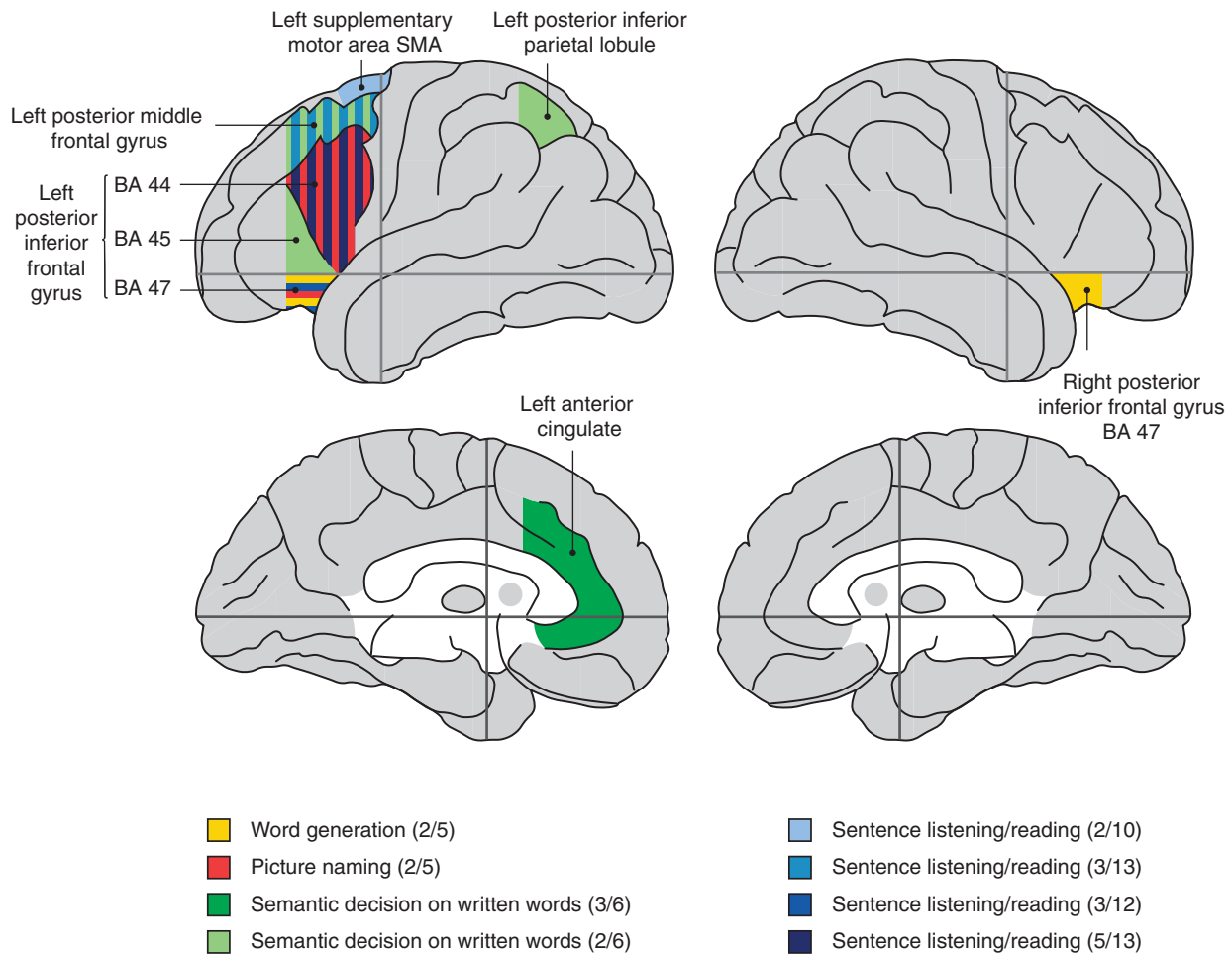
Figure 1 Schematic representation of brain areas reliably reported to be more strongly activated in L2 compared to L1 (based on a meta-analysis by Indefrey). Colors indicate the task in which stronger L2 hemodynamic activation has been found. Numbers in brackets refer to the number of studies reporting a brain area and the number of studies that examined this area. SMA, supplementary motor area.

all processing components involved in reading as well as semantic decision. Stronger activation for semantic decisions on L1 compared to L2 words has been reported for the anterior left middle temporal gyrus in late L2 learners with moderate to high proficiency. Stronger activation for semantic decisions on L2 compared to L1 words has been reported for the left posterior middle frontal gyrus, the left posterior inferior frontal gyrus (Broca's area (BA) 45), the left posterior inferior parietal lobule, and the anterior cingulate gyrus (Figure 1). The areas tend to be found in bilingual speakers with lower L2 proficiency, whereas there is no obvious influence of L2 onset or use. As in word production, areas of potentially stronger activation in L2 word comprehension are not L2 specific but are also observed in the native language. They are part of a modality-independent semantic system found to be activated during semantic decisions on both words and pictures. Anterior left middle temporal lobe activation, by contrast, has

been found for semantic decisions on written words only, not on pictures, suggesting that a stronger activation of this area for L1 compared to L2 words may be due to L1-specific word representations at the graphemic or lemma level.

The processing of inflectional morphology in a L2 has been investigated with fMRI. Japanese learners of English were visually presented with English verb stems and had to choose between correct ('caught' and 'smelled') and incorrect past tense forms ('catched' and 'smold'). After a 2-month training involving explicit instruction on verb inflection, 13-year-old beginners of English showed activation of the dorsal left posterior inferior frontal gyrus (BA 44/45) when performing the task. The activation was stronger the better participants performed on the task. The area overlapped with the activation area observed when the same task was conducted with Japanese verbs. In a follow-up study, the same task was used on 19-year-old learners of English, who by this age

had 6 years of English instruction. Again, increased activation of left posterior inferior frontal gyrus (together with parieto-occipital and motor areas) was found when comparing the past tense task to a control task. A subgroup of high-performing advanced learners showed much weaker activation in the inferior frontal gyrus area, and the activation was also weaker in older compared to younger participants when performing the task with Japanese verbs. Taken together, the two studies suggest a change in the linkage between performance on a certain linguistic task and accompanying hemodynamic responses. The activation level of an area may initially be positively correlated with performance but show a negative correlation after years of practice. The positive correlation may simply reflect neural and behavioral consequences of the same underlying factor – the degree of effort put into the task. By contrast, a negative correlation may be explained by assuming that neural structures support the linguistic process involved in the task more effectively, which means that they must have in some way adapted to the process. Note, however, that 'adaptation' is not the only option to interpret reduced activation with better performance. Alternatively, a different but more effective brain region might have taken over or such a negative correlation might occur because the task is performed in a different, more effective way that no longer involves the original process.

Sentence-Level Processing

A number of studies have presented proficient bilingual speakers with spoken or written sentences. Although some regions have been reported as being more strongly activated in L1 compared to L2 sentence or story comprehension, no region has been reported in more than one study. Regions that have been reliably replicated as being more active in L2 than in L1 sentence comprehension are mainly found in the left posterior frontal gyrus (middle frontal gyrus, BA 44 and BA 47 of the inferior frontal gyrus, and the supplementary motor area; Figure 1). Stronger L2 activation of these regions has typically not been observed when subjects simply read or listened to sentences or stories, but it seems to depend on the use of additional judgment tasks. In particular, stronger L2 activation of BA 44 has been found consistently when bilinguals were asked to distinguish between correct sentences and syntactic violations. Another shared, but not distinctive, feature of the studies reporting stronger L2 activations seems to be late L2 onset (after age 10 years). In summary, the available hemodynamic studies suggest that proficient L2 speakers recruit the same cortical areas (but

possibly stronger) for L2 as for L1 sentence-level comprehension.

We investigated changes in Broca's area activation in beginning Chinese learners of Dutch as an L2 for more than 1 year. Significantly stronger activation of Broca's area for listening to sentences compared to word sequences first emerged after 6 months of learning the new language, when the participants still scored in the low to moderate range of a standardized Dutch proficiency test. These findings suggest that brain regions involved in L1 syntactic processing are relatively soon also recruited for the processing of a new language.

Recruitment of similar brain areas for L1 and L2 sentence processing has also been observed in sentence production. Comparable activation levels for the two languages in frontal and temporal regions were found when bilingual speakers were asked to covertly report events of the previous day in their first and their L2s or when they generated sentences in a much better controlled task. Note, however, that in these studies differences between L1 and L2 in the exact localization of activation peaks in the frontal cortex (not in temporal cortex) were found, and these localization differences covaried with the age of L2 acquisition.

Electrophysiological Correlates of L2 Acquisition

The general picture that emerged from this overview of hemodynamic studies is that L2 processing involves the same brain structures as L1 processing but may be associated with greater activity. Complementary evidence for possible underlying processing differences may be provided in studies using electrophysiological measures such as the event-related potential (ERP) which are sensitive to the time course of neuronal activity. Past L2 research has mainly characterized L2 responses of speakers with preexisting L2 knowledge, whereas recent work has examined L2 learning.

Sublexical Processing

Studies of sublexical processing in bilinguals and learners have focused on the mismatch negativity response (MMN) observed in a passive listening task in which a sequence of standard sounds (e.g., syllables) are presented with an occasional mismatching sound. An MMN is taken as evidence that the brain has detected a difference between the two sounds. This technique has been used to examine whether native speakers and L2 learners show different electrophysiological responses to features of phonemic sounds that are distinctive in an L2.

Several studies have shown that an MMN response can be observed in both younger and older L2 learners, but that L2 learners who start relatively late show reduced amplitude MMN for L2 phonological contrasts.

Word-Level Processing

Several studies have compared L1 and L2 N400 responses to single words. The N400 is a negative brain potential deflection, peaking approximately 400 ms after stimulus onset, that is modulated by associative, semantic, or meaning-identity relations between a word and its context. Similar patterns of both associative and semantic priming have been shown in L2 N400 responses compared to L1 responses, but the pattern is modulated by L2 proficiency, with lower proficiency being associated with reduced conceptual but not associative priming. Current studies suggest that priming effects operate similarly in L1 and L2 single-word recognition, with some modulation by overall L2 proficiency.

Sentence-Level Processing

Although the priming effects that have been observed in L2 comprehenders suggest similar word-level semantic processing in L1 and L2, this does not necessarily mean that the semantic integration of words into a sentence context works equally well. Studies of L2 responses to semantic violations in sentence contexts have observed the N400 response to L2 violations, and most have shown a similar topographical distribution of this effect for L2 violations compared to L1 violations, sometimes with a delayed peak in the N400 response. Studies that have directly contrasted age of acquisition and proficiency have found that both age of acquisition and vocabulary proficiency make independent contributions to the longer latency of the N400 response in L2 readers. There is some variability in L2 response patterns, however, because no difference in the N400 violation response or a greater N400 violation response have both been observed. In most studies, however, the L2 N400 response in sentence contexts has been observed with a similar topography.

Two other electrophysiological violation responses are elicited by grammatical rather than semantic violations in native speakers. One is a late positive deflection (P600, or syntactic positive shift), and the other is a left anterior negativity (LAN) that is observed between 200 and 500 ms after a grammatical violation. L2 ERP studies have usually compared grammatical violation responses versus controls. Studies examining responses to phrase structure violations have demonstrated the P600 in early starting and/or proficient L2 learners, but the LAN has only been shown for relative

high proficient L2 users, and some reports have also shown an atypical P600 in late L2 learners. Other studies have observed the P600 response for certain combinations of L1 and L2 pairs but not others. The existing patterns across studies are not easily reconciled, but it appears that L2 P600 effects are more likely to be observed if the L1 and L2 belong to the same language family (e.g., Western Germanic for German and Dutch), if one takes into account the proficiency of the speakers tested. The limited number of magnetoencephalographic L2 studies have revealed similar variability in responses to grammatical violations. Whereas some studies have shown relatively early (in time) L2 violation responses, others have not. Finally, comparisons of L1 and L2 responses to open- and closed-class words (without grammatical violations) have shown L1/L2 response differences for closed-class words in later-starting L2 learners even though similar responses to L1 and L2 open-class words have been shown.

Studies have examined violation responses in L2 learners as they acquire the L2. Similar to the results in learners with existing knowledge, P600 violation responses have been observed when learners have acquired L2 proficiency. Although no early (in time) components have been observed in learners, some studies have observed an N400 response to grammatical violations in L2 learners.

These results show that responses to grammatical violations in an L2 can be observed using electrophysiological techniques, but the exact form of the violation response varies greatly, more so than semantic effects. Two factors, age of acquisition and global language proficiency in the L2, appear to play an important role in the appearance of all of these responses.

Conclusions

Neurocognitive studies of bilingual language processing have indeed provided some answers to the question of how an L2 is implemented in the brain. First, one of the three hypothesis mentioned in the introduction, the 'radically different' hypothesis, can clearly be rejected. L2 processing does not use any new brain areas nor does it give rise to new electrophysiological effects that are not found in L1 processing. With respect to the other two hypotheses, the situation is more complex. Whether L2 processing uses the same processes and brain structures as L1 processing or just a subset seems to depend on the processing level, and hemodynamic and electrophysiological data are not always in agreement.

The available evidence for sublexical processing suggests no fundamental differences. Some nonnative phonemic contrasts can be acquired to a level at

which MMN responses and hemodynamic responses are indistinguishable from those observed for native phonemic contrasts. Insofar as nonnative contrasts elicited differential responses, these seem to be of a quantitative rather than a qualitative nature, such as a reduced MMN or a stronger hemodynamic activation of the left inferior frontal gyrus, which is also active for phonemic decisions in the native language.

Similarly, at the lexical level, electrophysiological and hemodynamic data are in agreement that there are no L1/L2 differences with respect to the type or location of responses. Brain responses may be modulated by proficiency in the L2, with lower proficiency being correlated with reduced electrophysiological semantic priming effects and stronger recruitment of primarily left posterior frontal regions. Note, however, that there is hemodynamic evidence for L1-specific processing of written words. It is also important to note that the overall observation of similar recruitment of brain areas for L1 and L2 lexical processing may only hold at the level of larger areas, but within those areas, L1 and L2 may be separated to some extent in every individual bilingual.

Although below the sentence level the available data generally support the 'more of the same' hypothesis, electrophysiological and hemodynamic data give different answers with regard to sentence processing. Whereas hemodynamic studies report similar L1 and L2 processing areas with age of acquisition-dependent modulations (stronger recruitment of Broca's area in comprehension and location of activation peak differences within Broca's area in production), electrophysiological studies suggest that one typical native language response to grammatical violations, the LAN, is less commonly observed in the L2. Also, it cannot be said that speakers who acquired an L2 early across the board show the P600 in the same way as native speakers (as might be expected from similar L1 and L2 hemodynamic activation patterns in early L2 learners). Whether or not a P600 is observed seems to depend on factors such as distance between languages and the exact type of violation in ways that remain to be more closely examined in the future. These electrophysiological data are the main source of evidence supporting the hypothesis that L2 processing may only use a subset of the processes that are engaged for the L1. Insofar as this holds true, L1/L2 differences seem to be restricted to the sentence level, most likely syntactic processing.

How, then, can it be that such differences do not seem to affect the set of brain regions that are hemodynamically activated in sentence processing? First, there are design differences that reduce the comparability of hemodynamic and electrophysiological studies. Hemodynamic studies used mainly correct sentences and within-subject L1/L2 comparisons, whereas ERP studies used grammatically violating sentences and L1/L2 differences were observed in group comparisons. Second, both the L1/L2 differences observed in ERP studies (scarcity of LAN or P600 in the L2) and the L1/L2 similarities in hemodynamic studies are negative findings, so it remains possible that evidence to the contrary will be found in the future. Third, and probably most important, the presence or absence of ERP components does not hinge only on certain neural populations being active or not but, rather, on their synchronous activity. It could be that what appears as a qualitative difference in the data is, in fact, a quantitative difference in the timing of neuronal activation. If L2 speakers were more variable with respect to the exact timing of the cognitive operations underlying, for example, the LAN response in native speakers, this ERP response might not be observable anymore, whereas the corresponding hemodynamic effect of neuronal activation might still be observable due to the much coarser temporal resolution of fMRI or PET.

In summary, the most parsimonious interpretation of the available neurocognitive data on L2 processing is a 'more of the same' account assuming that L2s are implemented in the same brain structures as the native language, but with quantitative differences in timing and strength of activation. Alternative views assuming qualitative L1/L2 differences, such as the 'subset' hypothesis, have some, although not conclusive, support from electrophysiological data but are not supported by hemodynamic data.

See also: Bilingualism; Language Development; Language: Cortical Processes; Language: Aphasia; Psycholinguistics; Sentence Production; Sentence Comprehension; Statistical Learning of Language; Word Production; Word Recognition; Word Learning.

Further Reading

Golestani N, Alario F-X, Meriaux S, LeBihan D, Dehaene S, and Pallier C (2006) Syntax production in bilinguals. *Neuropsychologia* 44: 1029–1040.

Gullberg M and Indefrey P (2006) *The Cognitive Neuroscience of Second Language Acquisition.* Malden, MA: Blackwell.

Hahne A (2001) What's different in second-language processing? Evidence from event-related brain potentials. *Journal of Psycholinguistic Research* 30: 251–266.

Indefrey P (2006) A meta-analysis of hemodynamic studies on first and second language processing: Which suggested differences can we trust and what do they mean? *Language Learning* 56: 279–304.

Lucas TH, McKhann GM, and Ojemann GA (2004) Functional separation of languages in the bilingual brain: A comparison of electrical stimulation language mapping in 25 bilingual patients

and 117 monolingual control patients. *Journal of Neurosurgery* 101: 449–457.

Osterhout L, McLaughlin J, Kim A, Greenwald R, and Inoue K (2004) Sentences in the brain: Event-related potentials as real-time reflections of sentence comprehension and language learning. In: Carreiras M and Clifton C Jr. (eds.) *The On-line Study of Sentence Comprehension: Eyetracking, ERP, and Beyond*, pp. 271–308. New York: Psychology Press.

Paradis M (2004) *A Neurolinguistic Theory of Bilingualism.* Amsterdam: Benjamins.

Perani D and Abutalebi J (2005) The neural basis of first and second language processing. *Current Opinion in Neurobiology* 15: 202–206.

Rossi S, Gugler MF, Friederici AD, and Hahne A (2007) The impact of proficiency on syntactic second-language processing of German and Italian: Evidence from event-related potentials. *Journal of Cognitive Neuroscience* 18: 2030–2048.

Stowe LA and Sabourin LL (2006) Imaging the processing of a second language: Effects of maturation and proficiency on the neural processes involved. *International Review of Applied Linguistics* 43: 327–351.

Tatsuno Y and Sakai KL (2005) Language-related activations in the left prefrontal regions are differentially modulated by age, proficiency, and task demands. *Journal of Neuroscience* 25: 1637–1644.

Weber-Fox CM and Neville HJ (1996) Maturational constraints on functional specializations for language processing: ERP and behavioral evidence in bilingual speakers. *Journal of Cognitive Neuroscience* 8: 231–256.