Second Language Research

http://slr.sagepub.com

Temporal dynamics of late second language acquisition: evidence from event-related brain potentials

Karsten Steinhauer, Erin J. White and John E. Drury

Second Language Research 2009; 25; 13
DOI: 10.1177/0267658308098995

Second Language Research 25,1 (2009); pp. 13–41

Temporal dynamics of late second language acquisition: evidence from event-related brain potentials

Karsten Steinhauer, Erin J. White and **John E. Drury**
McGill University

Received August 2007; revised January 2008; accepted February 2008

The ways in which age of acquisition (AoA) may affect (morpho)syntax in second language acquisition (SLA) are discussed. We suggest that event-related brain potentials (ERPs) provide an appropriate online measure to test some such effects. ERP findings of the past decade are reviewed with a focus on recent and ongoing research. It is concluded that, in contrast to previous suggestions, there is little evidence for a strict critical period in the domain of late acquired second language (L2) morphosyntax. As illustrated by data from our lab and others, proficiency rather than AoA seems to predict brain activity patterns in L2 processing, including native-like activity at very high levels of proficiency. Further, a strict distinction between linguistic structures that late L2 learners can vs. cannot learn to process in a native-like manner (Clahsen and Felser, 2006a; 2006b) may not be warranted. Instead, morphosyntactic real-time processing in general seems to undergo dramatic, but systematic, changes with increasing proficiency levels. We describe the general dynamics of these changes (and the corresponding ERP components) and discuss how ERP research can advance our current understanding of SLA in general.

Keywords: late second language acquisition, event-related brain potentials, ERPs, critical period, L2 morphosyntax, LAN, P600

Address for correspondence: Karsten Steinhauer, McGill University, Centre for Research on Language, Mind and Brain, School of Communication Sciences and Disorders, Faculty of Medicine, 1266 Pine ssAvenue Wesst (Beatty Hall), Montreal, Quebec H3G-1A8, Canada; email: karsten.steinhauer@mcgill.ca

 10.1177/0267658308098995

I Introduction

The objective of this article is to elucidate the neurocognitive bases of late-acquired second language (L2). It is motivated by recent debates of the 'critical period hypothesis' (e.g. White and Genesee, 1996; Birdsong, 1999; Hakuta *et al.*, 2003) and focuses primarily on grammar-related processes for which fundamental differences between first language (L1) and late-acquired L2 have been postulated (Bley-Vroman, 1989). Electrophysiological data will be reviewed and tentative answers will be offered to questions such as the following:

1) Does late-acquired L2 involve the same neurocognitive mechanisms as found in native speakers?
2) Do the neurocognitive substrates of late-acquired L2 processing change with increasing L2 proficiency?
3) Is there evidence for a critical period in the acquisition of L2 grammar?
4) What might studies of the acquisition of artificial languages reveal about natural language acquisition/processing?

The evolutionary uniqueness of human language – as well as the remarkable effortlessness with which children acquire full language competence despite impoverished input stimuli – has led to the suggestion of an innate language-specific learning capacity. 'Universal Grammar' (UG), for example, is believed to constrain the representation of language, limiting the kinds of grammars that children adopt. According to some researchers, access to UG principles may be lost after a critical period (ending around puberty) due to brain maturation.[1] In comparison to young/early L2 learners, adult/late L2 learners display selective problems in phonology (foreign accent) and morphosyntax (Flege *et al.*, 1999; White, 2003; Johnson and Newport, 1989), while lexical learning (Eubank and Gregg, 1999), lexical–conceptual processing (Weber-Fox and Neville, 1996; Hahne, 2001; Wartenburger *et al.*, 2003), and certain aspects of syntax (White, 2003) often are relatively intact. According to the 'fundamental difference

[1] For the purpose of this article the concept of a 'critical period' will include the notion of a sensitive period (with a more gradual offset) (see Harley and Wang, 1997; Hyltenstam and Abrahamsson, 2003 for a discussion of the similarities and differences between these definitions). Moreover, this definition does not preclude the possibility that various linguistic subdomains (such as phonology, morphology, syntax, conceptual and logical semantics) may each be subject to distinct critical periods (e.g. Long, 1990). Unless stated otherwise, this article focuses exclusively on syntax and morphosyntax.

hypothesis' (Bley-Vroman, 1989), young children acquire their L1 implicitly with UG mechanisms, whereas late language learners depend largely on explicit, domain general cognitive functions (see also DeKeyser and Larson-Hall, 2002; DeKeyser, 2003; 2005).

While most researchers agree on the existence of critical/sensitive periods in L1, these notions are controversial in late L2 acquisition and processing. First, it is not clear whether late L2 learners can reach native-like competence. Whereas some studies suggest that later Age of Acquisition (AoA) of the target language virtually eliminates native-like attainment (Johnson and Newport, 1989), others suggest that such attainment is possible (White and Genesee, 1996; Birdsong, 1999).

Second, whereas some studies have reported the predicted discontinuity of AoA effects after the postulated critical period (Johnson and Newport, 1989), others have observed a monotonous AoA function incompatible with the claim of a critical period for language learning (Flege, 1999; Birdsong and Molis, 2001; Hakuta *et al.*, 2003). To date, the strongest evidence regarding a critical period has been provided by phonological learning (particularly phoneme discrimination) rather than grammar learning (Flege, 1999). Interestingly, one of the strongest pieces of evidence in favour of a critical period in L2 grammar learning compatible with fundamental difference hypothesis – namely that late L2 learners stabilize at some point short of native-like grammar attainment (so-called 'fossilization') – has most recently also been discussed in terms of phonological/prosodic interference from L1 (Goad and White, 2004; 2006).

Third, AoA effects have been explained in a variety of ways (Birdsong, 1999), including the loss of language-specific learning mechanisms (Bley-Vroman, 1989; Pinker and Prince, 1994), the advantage of small working-memory capacities in childhood (the 'less is more' hypothesis; see Newport, 1993), and 'neural commitment' or 'entrenchment' and consequent interference of L2 by earlier-learned knowledge (L1; Marchman, 1993). Given the somewhat unclear situation and lack of appropriate experimental techniques, researchers in the field of second language acquisition (SLA) have recently suggested that cognitive neuroscience may provide fruitful answers to these issues (Doughty and Long, 2003; Eubank and Gregg, 1999). This is reasonable given that both the critical period hypothesis and the fundamental difference hypothesis are essentially neurobiological hypotheses, and their proponents generally imply claims about neuronal differences in language representation between L1 and (at least) late acquired L2.

Age of acquisition and critical period effects do not seem to affect all aspects of SLA to the same extent and can be distinguished along several dimensions. First, they affect final attainment in L2 rather than rate of learning (e.g. DeKeyser, 2005; Krashen *et al.*, 1979). Second, they concern certain linguistic sub-domains (phonology and morphosyntax) more than others (lexical learning, semantic integration). Most recently, Clahsen and Felser (2006a; 2006b) have suggested a more fine-grained distinction for age effects on grammar processing. Their 'shallow structure hypothesis' claims that whereas it should be possible for late L2 learners to process morphology in a native-like manner, this should not hold for (non-local aspects of) syntactic processing.

In what follows, we first review the relevance of work using event-related brain potentials (ERPs) to study language processing in both L1 and L2 (Section II). In Section III we review evidence suggesting that:

- Late L2 learners can, in fact, be shown to display native-like ERP patterns; and
- the evidence in favour of the existence of principled limits on the range of L2 learnable properties of natural language – e.g. Clahsen and Felser's (2006a; 2006b) shallow structure hypothesis – are currently without compelling empirical support.

Finally, in Section IV we outline a hypothetical trajectory for L2 learning as indexed by the temporal dynamics of the emergence of various different ERP responses and suggest some directions for future research.

II Event-related brain potentials

Event-related brain potentials (ERPs) reflect the real-time electrophysiological brain dynamics of cognitive processes with an excellent time resolution in the range of milliseconds. Distinct ERP components (waveforms with either positive or negative polarity) have been identified for syntactic processing (Steinhauer and Connolly, 2008). Difficulties in semantic processing typically elicit centroparietal negativities that peak about 400 milliseconds post stimulus ('N400s'; Kutas and Hillyard, 1980), and depend particularly upon bilateral temporal lobe structures (Halgren *et al.*, 2002; Van Petten and Luka, 2006). Syntactic processes have been linked to two ERP components: an early (150 ms–500 ms) left anterior negativity (LAN) and a late (600 ms–900 ms) centroparietal positive shift (P600). Violations of phrase structure (Neville *et al.*, 1991; Weber-Fox and Neville, 1996; Hahne, 2001) and of morphosyntactic constraints (Friederici *et al.*, 1993; Coulson *et al.*, 1998; Gunter *et al.*, 2000) often yield LANs – which have been hypothesized to be linked to rule-based automatic parsing (Friederici *et al.*, 1999; Hahne and Friederici, 1999).[2] Neural generators of the LAN have been identified in the (left) prefrontal cortex, notably Broca's area (see Knoesche *et al.*, 2000; Friederici, 2002). Although named for its distribution across the scalp, some variability in the topography of LAN components has been reported, including both bilateral anterior (i.e. Hahne and Jescheniak, 2001; Hagoort *et al.*, 2003a) and more left temporal distributions (Neville *et al.*, 1991). Syntactic processing difficulties also elicit P600s, taken to reflect the costs of controlled syntactic processing (e.g. for re-analyses, see Osterhout and Holcomb, 1992; Friederici *et al.*, 1999; Friederici, 2002; or, for integration, see Kaan *et al.*, 2000; Phillips *et al.*, 2005). Grammar violations often elicit a biphasic ERP pattern consisting of a LAN followed by a P600 (Steinhauer and Connolly, 2008).

If indeed LANs are generally linked to automatic, implicit grammar processing, one would predict that these early ERP components in particular should be the most difficult to elicit in late L2 learners (i.e. more likely to be constrained by AoA or critical periods than, for example, the P600). As will be demonstrated below, although this prediction has been confirmed by a number of early ERP studies, more recent work reveals that the relationship between AoA and ERPs is more complex.

ERP studies on L2 acquisition have generally adopted the 'violation paradigm' used in L1 research: grammatical violations that become evident at a specific target word are directly compared with matched control sentences that do not contain a violation. Grammaticality judgements at the end of a sentence serve as the standard task during

[2] It is important to note (we will also return to emphasize this point below) that the neurocognitive etiology of ERP components is by no means a resolved matter. For our purposes here we will maintain/adopt the view that LAN effects reflect automatic aspects of sentence processing, and although we recognize that this assumption is debatable, further discussion of other possible interpretations of these effects is beyond the scope of this present article. For some recent reviews and other discussions of LAN-type ERP effects in sentence processing, see Martin-Loeches *et al.*, 2005; Steinhauer and Connolly, 2008.

the experiment. ERP studies can be broadly divided into three kinds of experimental design:

- group studies contrasting L2 learners with native speakers or different groups of L2 learners;
- longitudinal studies testing changes within participants while they acquire the target language; and
- paradigms using artificial miniature languages that can be learned to high proficiency within a short period, allowing one to cover large ranges of proficiency changes within the same participants and at the same time control for many potential confounds.

Like the behavioural studies mentioned above (e.g. Johnson and Newport, 1989), early ERP studies argued in favour of the critical period hypothesis (i.e. a late AoA results in fundamental differences in the way syntactic information is processed; Weber-Fox and Neville, 1996; Hahne, 2001). Weber-Fox and Neville explored whether the neurocognitive processes underlying semantic and syntactic processing in bilinguals are impacted by AoA. They found that when presented with semantically anomalous sentences, all of the bilinguals – regardless of AoA – elicited an N400 response, although this effect showed a delayed peak latency in those who learned English after age 11. This suggests that semantic processing strategies rely on largely identical neurocognitive mechanisms in both one's L1 and L2, but that for late L2 learners, lexical retrieval and/or semantic integration may be somewhat slower.

In contrast, late L2 acquisition had a much more pronounced effect on syntactic processing and was associated with changes in the distribution, amplitude and latency of the LAN and P600 components. In response to syntactic phrase structure violations, the early L2 learners (AoA < 11 years) elicited a left-lateralized negativity similar to that of native speakers. For the late learners (11–13 and 16+ years), however, this negativity was bilaterally distributed, and for the oldest group it was actually greater over the right hemisphere than the left. Thus, although all groups elicited a negativity in this time window, only the early bilinguals elicited a LAN similar to thsat of native speakers. Similarly, the P600 elicited by the early learners (AoA < 11 years) was identical to that of native speakers, was delayed in those who acquired English between

11–13 years, and was not present at all for the oldest group of L2 learners.

Together, this and subsequent work (Hahne and Friederici, 2001) suggests that while the neurocognitive basis of semantic processing may be relatively unaffected by AoA, the neural mechanisms underlying syntactic processes appear to be sensitive to delays in L2 acquisition. The fact that age-related differences were observed in both the LAN and P600 responses (indexing early syntactic parsing and subsequent syntactic re-analysis/repair processes, respectively) provided powerful evidence for the claim that syntactic processing relies on fundamentally different neural-cognitive structures in late L2 learners compared to native speakers and early L2 learners.

III Proficiency and native-like attainment

The case may have been closed as to the fundamental role of AoA in restricting the recruitment of neurocognitive mechanisms necessary for native-like grammatical processing in L2, if it were not for one important confound: the bilingual's level of proficiency. For example, AoA was negatively correlated with L2 proficiency in Weber-Fox and Neville (1996) as measured by standardized tests of English grammar, self-reported proficiency and acceptability judgement accuracy. It is thus impossible to determine whether the lack of LAN/P600 effects in late learners was due to AoA or proficiency. The same difficulty holds for Hahne and Friederici (2001) who found no differences between violation and control conditions but who did not report proficiency levels among the late Japanese–German bilinguals they tested.

In fact, evidence from neuroimaging (PET/fMRI) studies clearly reveals the role of L2 proficiency in influencing the recruitment of native-like neurocognitive mechanisms (see Perani *et al.*, 1998; Abutalebi *et al.*, 2001; Wartenburger *et al.*, 2003). In Perani and colleagues' (1998) imaging study, it was shown that at relatively low levels of L2 proficiency, whereas in learners with very high levels of L2 proficiency they had acquired their L2 early or after puberty. However, these findings were relatively unspecific in terms of which psycholinguistic sub-processes were actually reflected by proficiency-dependent brain activation, and it is important to note that fMRI (and PET) lack the temporal resolution to detect fine-grained differences in the rapid dynamics of language processing. As canvassed above (Section II), up until 2001 there had been no ERP study similar to Perani *et al.*'s neuroimaging studies. Although the available ERP evidence at that time generally supported the notion of a critical period for syntactic (but not semantic) processing, in all of the relevant studies second language learners were not even close to the proficiency level of native speakers.

Friederici *et al.* (2002) presented the first ERP study directly addressing the AoA/proficiency confound. They trained adult (post-critical-period) participants in an artificial miniature language (BROCANTO) to a native-like level of proficiency in both production and perception (95% accuracy criterion). Unlike artificial grammars in previous studies (Reber, 1967; Baldwin and Kutas, 1997; Reber and Squire, 1999), BROCANTO both conformed to UG requirements and could be used to actually communicate complex propositions. Sentences referred to the moves of a complex (chess-like) computer-based board game for two players (Figures 1a and 1b). For example, the Brocanto sentence *aak füne plox prez nöri aaf trul* translates into 'the round Plox-piece horizontally captures the Trul-piece'. Participants practised the language BROCANTO by verbally communicating their moves while playing against each other at separate computer monitors. These practice sessions were distributed

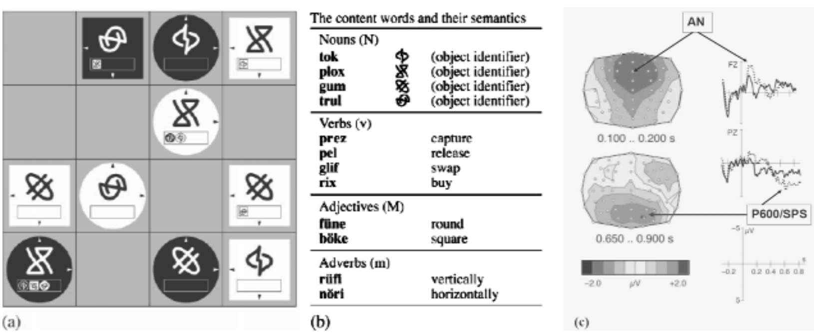

Figure 1 Initial evidence for proficiency-dependent ERP's in late L2 grammar: auditory study on artificial language learning by Friederici *et al.*, 2002; (a) chess- like 'Brocanto' game board used to engage subjects in speaking BROCANTO; (b) vocabolary; and (c) voltage maps and ERP plots illustrating the main findings of the artificial language experiment

Notes: After training tovery high ('native-like') proficiency level, subjects displayed the typical 'native-like ERP patterns of syntactic processing: an early anterior negatively (AN) followed by a P600. (Adapted from Friederici *et al.*, 2002: Figures 2 abd.)

across several days and lasted up to 25 hours in total until participants reached criterion proficiency in both production and comprehension.

Intriguingly, in the subsequent ERP experiment, syntactic word category violations in BROCANTO elicited early anterior negativities (followed by a P600 component) (Figure 1c) that had been taken as a hallmark of native-like sentence processing in studies of natural language. This ERP finding was related to high performance in both a visual probe verification task (89% correct) and a grammaticality judgement task (93%). In contrast, a control group that had received only vocabulary but no grammar training performed well only in the probe detection task (86%) but not in the grammaticality judgement task (58% correct). As expected, this control group did not display any ERP differences between grammatical and ungrammatical sentences.

This study raised a number of important new questions: Can qualitative neurocognitive changes during language learning be monitored within participants using ERPs? What are the limitations of SLA models based on artificial language paradigms? A number of recent studies have provided some preliminary answers to these questions.

1 ERP studies showing native-like patterns in L2

Another key question raised by the BROCANTO findings is whether similar effects of native-like ERP patterns in late L2 acquisition can be replicated in natural language learners. An ERP study conducted in our own lab with visually presented stimuli (Steinhauer *et al.*, 2006) examined late French and Chinese learners of English at two different proficiency levels (high vs. low) with respect to their processing of syntactic word category violations and contrasted them to native English speakers. All L2 learners acquired English after the age of 12 years and would be categorized as late learners (Birdsong, 1999), but the mean AoA was in fact higher (Chinese learners: 15.0 years, range:12–28 years; French learners: 19.0 years, range: 15–23 years). Mean age during the experiment was 27.3 years (range: 19–36 years with no differences among groups). The high/low proficiency grouping was based on performance in a sentence completion test (cloze test), with ‘high proficiency’ requiring at least 90 per cent correct completions.

Given several severe problems with the stimulus materials of previous ERP syntax studies (for discussion, see Steinhauer and Connolly, 2008), a new syntactic violation paradigm was developed for the present study that counterbalanced the occurrence of target words – in italics in examples (1–4) below – and their context in violation conditions (2 and 4) and control conditions (1 and 3).

Sentence conditions used in the ERP experiment

1) The man hoped to enjoy the meal with friends.
2) The man hoped to **meal* the enjoy with friends.
3) The man made the *meal* to enjoy with friends.
4) The man made the **enjoy* to meal with friends.

The ERP pattern for the native speakers consisted, as predicted, of a LAN followed by a P600 (see Figure 2: left panel). These two components presumably reflected an early interruption of the relatively automatic parsing process (LAN) and subsequent re-analysis and repair of sentence structure (P600; Hahne and Friederici, 1999). In contrast, low proficiency French and Chinese learners elicited only a P600 without any indication of an early negativity (Figure 2: right panel). The absence of a LAN effect in this group suggests – in line with previous studies (e.g. Weber-Fox and Neville, 1996; Hahne, 2001) – that late learners with a moderate L2 proficiency were not able to automatically process the grammar violation in a native-like manner within the first 500 ms post violation.

However, high proficiency French and Chinese learners did elicit a biphasic LAN/P600 pattern that was statistically indistinguishable from that of the native speakers (Figure 2: middle panel). In other words, despite late AoA, and independent of typological similarity between L1 and L2 (i.e. French–English vs. Chinese–English), the late learners displayed native-like ERP patterns of brain activation if they had reached a high level of proficiency.

Importantly, the same pattern found for French/Chinese learners of English was also found in a reading study on English-speaking learners of Spanish, showing that the effects of proficiency on ERPs were not specific to the L2 target language. Bowden *et al.* (2007) contrasted Spanish native speakers with two groups of late L2 learners at low and high levels of proficiency. The Spanish sentence materials included a semantic anomaly condition as well as a word category violation (modelled after the English paradigm in examples 1–4 above). The semantic anomaly

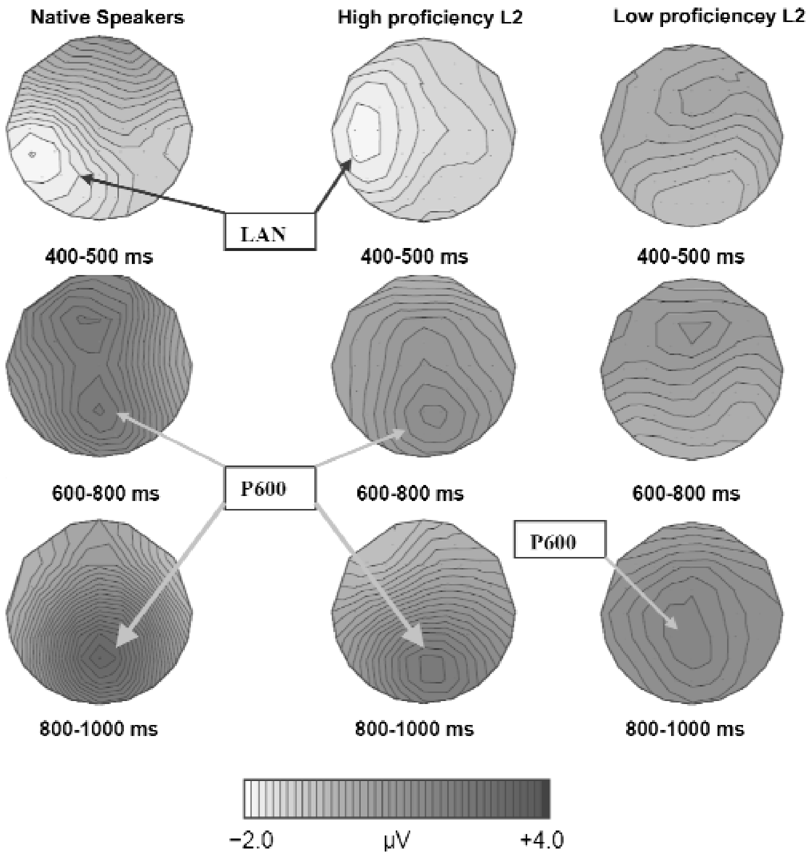

Figure 2 Voltage maps for the three groups in Steinhauer *et al.*, 2006

Notes: Natives speakers and high proficiency late L2 learners (collapsing across Chines/French L1) elicited statistically indistinguishable biphasic LAN/P600 patterns in response to syntactic violations. In contrast, low proficiency late L2 learners yielded only a P600.

yielded N400s in all three groups as expected. Syntactic violations elicited only small P600s (and no LANs) in low L2 proficiency participants, but native-like LAN/P600 patterns in high proficiency L2 learners.

Taken together, these studies suggest that the absence of LANs is, indeed, a typical pattern for late L2 learners (in line with Weber-Fox and Neville, 1996; Hahne, 2001; Hahne and Friederici, 2001). However, current data go beyond these studies by demonstrating that this pattern holds only for less proficient L2 learners. At higher levels of proficiency, L2 learners' ERPs consist of a biphasic sequence of components that is indistinguishable from that in native speakers.

The presence of LAN-type effects for non-native speakers in these studies turns out not to be limited to isolated findings. In fact, native-like patterns, including LAN-type effects, have also been reported for (morpho)syntactic aspects of language processing in virtually all recent L2 ERP studies from other labs that have looked at proficient late learners. For example, Kubota *et al.* (2004) found MEG evidence for early LAN-like magnetic fields for syntax violations in noun phrase raising constructions (*The man was believed /to have been killed/*was killed*). Ojima *et al.* (2005) studied conceptual–semantic anomalies and subject–verb agreement violations in both native speakers and low and high proficiency Japanese learners of English. For semantic anomalies, an N400 was observed in all groups. For morphosyntactic violations (subject–verb agreement), the authors found the LAN/P600 pattern typical of native speakers, just a LAN without a P600 for highly proficient L2 learners,[3] and no effects at all for low proficiency L2 learners. Rossi *et al.* (2006) used an auditory ERP paradigm to investigate adult German learners of Italian and adult Italian learners of German with respect to:

1) syntactic word category violations;
2) morphosyntactic agreement violations; and
3) double violations combining (1) and (2).

In both German and Italian, the high proficiency late learners displayed the same ERP pattern as native speakers, consisting of early anterior negativities and subsequent P600s. Strikingly, this held for all three conditions. At low proficiency, the L2 learners displayed smaller and delayed P600s, and the LAN for agreement violations was absent.

In summary, the foregoing findings strongly suggest that it is possible for L2 learners to elicit native-like ERP patterns. However, there are

[3] The lack of a P600 in this case requires further scrutiny, raising issues beyond the scope of this article. Ojima *et al*. point out that the positivities may have simply occurred later than the time windows that they examined (i.e. later than 850 ms post stimulus). However, they also make the intriguing (and testable) suggestion that the P600 may not have been found in these learners because of the lack of subject–verb agreement in Japanese, the participants' L1. This would suggest that a prerequisite for eliciting native-like ERP responses is the ability to transfer linguistic processing strategies directly from one's L1; this is a claim that warrants further investigation.

several important further points to be made about proficiency effects in ERPs. First, such effects are not limited to second language learning. A recent study by Pakulak and colleagues (2004) shows that even within L1, higher proficiency levels result in stronger left lateralization of LAN effects and larger amplitude P600s (for related findings, see also Weber-Fox *et al.*, 2003).

Second, even within the same L2 learners, relative L2 proficiency levels may differ between specific structures resulting in distinct ERP patterns. Hahne *et al.* (2006) presented late Russian–German bilinguals with two kinds of L2 inflectional violations involving past participles and noun plurals. Behavioural data showed these bilinguals were relatively more proficient with the past participles; this is a finding that was also apparent in the ERP data. The bilinguals showed P600 effects for both violations but a LAN effect only in response to past participle violations. Native speakers, in contrast, exhibited the biphasic response for both types of violation. This suggests that high levels of proficiency with specific structures are a prerequisite for automatic grammatical processing as indexed by the LAN. The specific difficulties a given L2 learner encounters may partly be influenced by transfer effects from the respective L1 (see also below).

Third, it may be the case that the mode of language learning/training may modulate ERP effects for L2 processing even where relative proficiency does not differ. Morgan-Short *et al.* (2007) employed a modified version of the artificial language used in Friederici *et al.* (2002) (BROCANTO 2) in a study in which participants were assigned to one of two groups: an 'explicit' training group that received classroom like instruction or an 'implicit' group that was exposed to a large number of sample sentences without formal instruction. Each participant was tested on both syntactic word category and agreement violations in two ERP sessions at two different levels of L2 proficiency (low and high). Importantly, implicit and explicit groups reached equally high levels of proficiency. However, the biphasic LAN/P600 effect was elicited, in both conditions, only in the implicit group. As these data suggest, distinct brain systems underlie implicit vs. explicit learning; they cast some doubt on the notion that late L2 learning generally depends upon explicit processing (for relevant discussion, see DeKeyser and Larson-Hall, 2002; DeKeyser, 2005).

2 Limits on native-like attainment in late L2 acquisition?

Clahsen and Felser (2006a; 2006b) (henceforth 'C&F'), as noted in our introduction, have suggested that although it may be possible for late L2 learners to process morphology in a native-like manner, this may not hold for syntactic processing (i.e. their 'shallow structure hypothesis'). More specifically, C&F propose that late L2 learners are unable to compute adequate hierarchical syntactic representations and, thus, are expected to rely largely on compensatory processing strategies involving semantic and pragmatic information. Processing problems are especially expected, on their view, for non-local structural properties, such as long-distance dependencies.

First, note that the general view that native-like L2 acquisition is restricted to morphological aspects of language cannot be maintained in the face of studies reviewed above; e.g. both the Steinhauer *et al.* (2006) and Rossi *et al.* (2006) findings for syntactic violations. Second, if the crucial dimension separating learnable from unlearnable aspects of grammar in L2 turns on the locality of the relevant linguistic dependencies, we must first be clear whether this refers to linear or hierarchical locality. Concerning locality in the hierarchical sense, although no previous ERP study has carefully examined the issue,[4] behavioural evidence strongly suggests that late L2 learners can indeed reach native-like proficiency levels with respect to constraints on long-distance dependencies (e.g. subjacency or empty category principle; see White

[4] Bahlmann *et al*. (2006) contrasted, within participants, the acquisition and processing of (1) a linear finite state grammar (FSG) with local dependencies [AB][AB][AB][AB]) and (2) a hierarchical phrase structure grammar (PSG) with long-distance dependencies (A[A[A[AB]B]B]B). In line with the large literature on non-linguistic artificial grammar learning – and unlike the miniature language studies using BROCANTO discussed in the main text – this ERP study investigated abstract sequences of syllables without semantic content. Accuracy in the grammaticality judgement task during the ERP session was very high in both PSG (96%) and FSG (98%). Violations generally elicited P600-like (or P300-like) positivities (400 ms–750 ms), the amplitude of which was position dependent for the PS but not for the FS grammar. Only FSG violations elicited additional posterior negativities (300 ms–400 ms) that were interpreted as reflection of conscious expectancies and explicit rule knowledge. A similar fMRI study had suggested involvement of distinct brain circuits including Broca's area in PSG and the frontal operculum in FSG (Friederici *et al.*, 2006). Due to several critical (but avoidable) confounds in the design (such as grammar type and number of relevant elements), the differences between conditions are difficult to interpret. However, to the extent that the study can be related to linguistic grammar learning and processing, it clearly reveals the ability of adult learners to acquire new long-distance dependencies, presumably even implicitly (given the absence of the posterior negativity in PSG violations). We believe these results are incompatible with what Clahsen and Felser's shallow structure hypothesis would predict.

Table 1 Critical Double Case Violation from Mueller *et al.* (2005)

Correct		
Ichi wa no kamo -GA	ni hiki no neko -O	tobikkeru tokoro desu
One [class][gen] duck [nom]	two [class][gen] cat [ACC]	jump over take place
Case violation		
Ichi wa no kamo -GA	ni hiki no neko *-GA	tobikkeru tokoro desu
One [class][gen] duck [nom]	two [class][gen] cat *[NOM]	jump over take place

Notes: [class] = classifier, [gen] = genitive, [nom] = nominative, [acc] = accusative

and Genesee, 1996). In the case of locality in the linear sense, as noted by C&F, recent ERP work by Mueller *et al.* (2005) is relevant. In this study, adult Japanese and German native speakers were tested using a miniature version of Japanese. ERPs were recorded for violations involving word category, case and classifiers, both before and after the Germans engaged in language training. The relevant condition for C&F's locality thesis is the case violation that involved a double occurrence of nominative case (Table 1).

These violations elicited a frontal negativity (referred to by Mueller *et al.* as an N400 effect) and a P600 in native speakers, but only a P600 in L2 learners after training. C&F take the absence of a negativity to indicate non-native levels of attainment in their late L2 learners. This would be consistent with their views, especially considering that the participants had been 'trained to a level of perfection at which they produced hardly any errors' (Clahsen and Felser, 2006b).

However, although it is true that these L2 learners did indeed attain an overall high level of proficiency, under closer examination their performance on the acceptability judgement task for the critical case violation was strikingly different from native speakers (84.1% vs. 98.4%; $p < .0001$) (recall our point above regarding the possibility of proficiency with specific structures within a given language). Moreover, in a follow-up study (Mueller, 2006), the same participants from Mueller *et al.* (2005) returned for further training and ERP testing. Strikingly, after this training the L2 learners' proficiency level increased significantly so that they were on par with native performance (95–99% acceptability judgement accuracy in both groups). And, at this point, the ERP effects for the double nominative violation yielded the biphasic N400/P600 in both groups showing that 'case information was used in the same time window as in native speakers.' This is clearly consistent with our understanding of the role of proficiency in attaining native-like ERP patterns and does not seem to fit with C&F's proposed constraints on L2 acquisition (i.e. the shallow structure hypothesis). However, although the presently available data point toward degree of proficiency rather than locality of dependency as being crucial in determining ERP profiles in late L2 learners, this is clearly an area in need of further study.

3 Longitudinal studies

All the studies on natural language learning reviewed in the foregoing section were group studies conducted with between-participant designs. It is well known that L2 learners exhibit larger inter-individual variability than native speakers on both behavioural performance measures and on measures of neural activity during language processing (Dehaene *et al.*, 1997), and these could clearly contribute to observed group differences. However, a few other studies have also demonstrated neurocognitive changes at early stages of SLA within the same participants in longitudinal within-participant designs, albeit with participants who attained only relatively low levels of proficiency (and thus did not elicit native-like brain responses). Nonetheless, as we discuss below, these findings are of interest in terms of what they may contribute to our understanding of neurocognitive changes during the early time course of L2 acquisition.

In a series of studies, Osterhout and colleagues investigated the early stages of both lexical-semantic and morphosyntactic knowledge acquisition in American learners of French during a 9-month introductory university course. Employing a priming paradigm, McLaughlin *et al.* (2004) found that lexical knowledge (i.e. the distinction between familiar words and legal pseudowords, and between semantically related vs. unrelated words) was reflected in N400 effects that became increasingly pronounced with higher levels of proficiency. The fact that ERP differences were observed in the absence of behavioural effects illustrates the greater sensitivity of ERPs as a measure of subtle changes in underlying processing mechanisms as compared to behavioural data. Osterhout *et al.* (2006) presented subject–verb agreement violations and correct control sentences (*Tu adores/*adorez le Français*) in a grammaticality judgement task. A subset of their participants (the 'fast' learners) exhibited an N400 to the ungrammatical verb structures after 1 month of instruction. This effect was replaced by a small P600 after 4 months, the amplitude of which subsequently increased with proficiency. Interestingly, no performance improvement or ERP effects were observed for number disagreement within noun phrases (*Tu manges des hamburgers/*hamburger pour dîner*), even at the end of the course (note this violation may have been less salient to learners since it is only orthographically and not phonologically realized). The relevant findings here are that:

- at very low proficiency, morphosyntactic violations may elicit N400s rather than LANs or P600s, but that increased proficiency results in qualitative changes in ERPs; and
- factors such as phonological recoding may play a role in developing native-like patterns of L2 syntactic processing.

Another longitudinal study has been conducted in our lab (White, Genesee, Drury, and Steinhauer, 2007). Korean students were tested before and after a 9-week intensive English course for adult learners at low-to-intermediate levels of proficiency. The critical conditions comprise:

- syntactic word category violations; and
- morphosyntactic agreement violations of tense and number.

Apart from early changes in another ERP component (the P200) reflecting improved reading skills, P600 components for morphosyntactic violations emerged by the second session. This is in line with the observations for French by Osterhout *et al.* (2006) discussed above.

In summary, ERPs are capable of reflecting both implicit and explicit aspects of processing of morphosyntactic information and can thus be used to distinguish between native-like and non-native on-line processing even before any such effects are evident in behavioural data. Whereas early ERP evidence initially suggested support for a critical period for SLA that prevents late L2 learners from relying on the same neurocognitive mechanisms as native speakers, these studies failed to take into account confounds such as AoA and proficiency. In contrast, more recent work that has disentangled these factors has provided strong evidence that, in fact, native-like processing of late acquired L2 is indeed possible in at least some domains of morphosyntactic processing. Further, we find that the currently available ERP data do not provide support for a principled division in L2 learnability between local and non-local types of linguistic dependencies (i.e. as required by the shallow structure hypothesis; C&F). Finally, longitudinal studies tracking the earliest stages of late L2 acquisition reveal suggestive systematicity to the transitions in ERP profiles that may track the path of L2 learning. We return to this in the next section.

IV The temporal dynamics of L2 acquisition

ERP patterns associated with language processing clearly change over the course of late L2 acquisition. In the earliest stages of learning, ERPs are insensitive to the distinctive properties of the L2. On the other end of the learning path, when learners have acquired high levels of proficiency, late L2 learners may show ERP responses that are indistinguishable from those of native speakers. Here we specify a hypothetical time course of changes in ERP patterns of late learners during L2 acquisition extracted from the wide range of results discussed earlier.

We suggest that with increasing proficiency, L2 learners' brain activation profiles typically approximate that of native speakers in a systematic way. While a number of variables (e.g. L1 background, learning environment, etc.) may modulate the starting point and exact time course of these transitions, we propose that a prototypical second language learner will be more likely to pass through these putative stages than not. The endpoint may also vary depending on the linguistic structure of interest, the target language, and the range of variability among native speakers of the target language (with respect to lateralization of LAN-type effects as a function of L1 proficiency, see Pakulak and Neville, 2004). Table 2 traces our hypothetical learning path from novice to native-like mastery of L2 (column 2) in terms of ERP components elicited by (morpho)syntactic violations (column 3), possible underlying cognitive processes indexed by such effects (column 4) and pointers to consistent relevant literature (column 5).

Two notes of qualification are in order. First, for the purpose of this exercise we assume idealized, highly motivated learners who are initially exposed to classroom instruction and then improve their language skills in an immersive environment with frequent use of the L2. Second, it is important

Table 2 ERPs reflecting different stages of morpho-syntactic proficiency in late L2

Stage	Proficiency level	ERP pattern	Cognitive processes	References
1	Novice	No difference	Indifferent perception; performance at or near chance level	Friederici, 2002; Mueller, *et al.*, 2005 (untrained controls); Rossi *et al.*, 2006 (2 conditions)
2	Very low proficiency	N400 or right-lateralized/posterior negativities	Difficulties during lexical access and integration; compensatory processing strategy, likely relyingon semantic plausibility and pragmatics	Weber-Fox and Neville, 1996; Hahne, 2001; Osterhout *et al.*, 2006 (1 month); White *et al.*, 2007 (regarding Broca's aphasics, see also Hagoort *et al.*, 2003b)
3	Low to intermediate	Small/delayed P600 (possibly preceded by N400s)	Beginning gram maticalization/ proceduralization	Osterhout *et al.*, 2006 (4 months); Steinhauer *et al.*,2006;
4	Intermediate	Larger/earlier P600	Late structural reanalysis/repair approaches native-like mecshnisms	Osterhout *et al.*, 2006 (9 months);
5	Intermediate to high/near native-like	Bilateral AN + P600	Near-native processing; early automatic + late controlled processing	Friederici 2002; Bowden *et al.*, 2007; Morgan-Short *et al.*, 2007
6	very high/ native-like	Lateralized LAN + P600	Native-like processing; early automatic + late controlled processing	Pakulak and Neville, 2004; Steinhauer *et al.*, 2006; Rossi *et al.*, 2006; Bowden, 2007 *et al.*; Morgan-Short, 2007 *et al.*

to stress that the etiologies of ERP effects seen in connection with language and other aspects of cognitive/perceptual processing are still a matter of considerable debate. Thus, what follows necessarily entails a number of assumptions that we acknowledge are oversimplified and/or likely to change.[5] Nonetheless, the overall picture laid out in Table 2 strikes us as:

[5] Two such assumptions/oversimplifications concern the LAN and P600 components as we have discussed them here (see footnote 2 regarding LAN effects). The P600s is very likely comprised of a number of distinct sub-components, each of which may reflect different kinds of cognitive processes; some of these may, moreover, be recruited across different cognitive domains (e.g. in music, mathematics, etc.; Steinhauer and Connolly, 2008).

- plausible;
- not without empirical support;
- implicit in many discussions of ERP effects and late L2 acquisition; and
- fairly easy to falsify.

Below (see Section IV.2) we offer some general considerations that we hope will contribute to the direction of future inquiry in this area of research.

1 ERPs in morphosyntax across the L2 learning span

- Stage 1: Initially, grammatical and ungrammatical structures cannot be distinguished, i.e. a violation is not recognized as such, and no differences in brain activity between violation and correct control conditions are observed. Note that the absence of effects is expected even if the learner is already familiar with the meaning of most content words but has no grammatical knowledge (e.g. the control group in Friederici *et al.*, 2002).
- Stage 2: At the next stage, violation conditions elicit N400s. This pattern would be expected, for example, if the violation occurs on a content word that the learner knows. Several cognitive processes may contribute to the N400 effect: the low probability for that word in that position; whole-form storage of (what will at later stages of higher proficiency be analysed as) morphologically complex forms, or perhaps explicit rule knowledge. The morphosyntactic violation is not yet recognized as such and so the anomaly is perceived as a lexical problem. The N400 effect observed by Osterhout *et al.* (2006) after one month of classroom instruction illustrates this stage. Complete reliance on declarative rule knowledge and compensatory strategies may result in this pattern. (Interestingly, a similar pattern for agrammatic aphasics who have to rely on alternative compensatory processing strategies is reported in Hagoort *et al.*, 2003b.) According to Ullman's Declarative/Procedural model (Ullman, 2001; 2005), N400s should be the standard brain response to any grammatical anomaly in beginning L2 learners in general, as the procedural/implicit system is initially not accessible.

- Stage 3: The third stage is characterized by the beginning of grammaticalization (Osterhout *et al.* 2006) or proceduralization (for discussion of possible mechanisms along these lines, see DeKeyser, 2005). The learner begins to identify the structural nature of the problem, and attempts to re-analyse or repair the problem, resulting in a (usually delayed) small P600. (Note though that it is unclear whether such effects reflect the same processes as in native speakers – see footnote 5). If negativities precede the P600, they resemble N400 rather than LAN components since it is hypothesized that at this stage the violation is still initially processed as a lexical problem.
- Stage 4: With increasing proficiency, the P600 amplitude increases (usually with a parietal maximum). This stage reflects systematic attempts to fix the structural problem in ways that are progressively more similar to native speakers. Thus, the P600 strongly resembles that of native speakers.
- Stage 5: If native speakers elicit LAN-like components preceding the P600, very proficient (near-native) L2 learners may elicit a frontal but less left-lateralized negativity followed by a P600 (e.g. Friederici *et al.*, 2002; see Figure 1 above). Note that bilateral frontal negativities for L1 have also been observed in native speakers who are less proficient in their L1 (Pakulak and Neville, 2004).
- Stage 6: At the highest L2 proficiency levels, no differences between native speakers and L2 learners are predicted. Note that the scalp distribution of LAN-like components displays some variability in native speakers, or may be absent altogether (see, for example, Osterhout and Mobley, 1995); however, the point here is that where these LAN effects reliably show up in natives, at this stage of L2 acquisition they should show up in non-natives as well.

It is important to emphasize that this sequence of ERP components (N400, [N400+] small P600, large P600, AN-P600, LAN-P600) could be expected to have different time courses and temporal dynamics for different morphosyntactic structures. In extreme cases, one structure may already display native-like ERP components while another one still elicits N400s or small P600s. The relative order in which structures reach higher levels of proficiency, we suggest, may be influenced by typological overlap with L1 (as well as individual differences in L2 exposure). For example, Chinese learners of English may have more problems than francophone learners with structures involving determiners, while the opposite pattern may be true for adverb placement. Unlike Clahsen and Felser (2006a; 2006b), we do not adopt strong assumptions regarding the notion that long distance dependencies should always be the most difficult structures.

2 Directions for future research

The various L2 studies discussed above can be distinguished along three dimensions, yielding an array of highly complementary experimental approaches:

1) within-participant longitudinal designs vs. between-participant designs comparing L2 learners with native speakers and/or other groups of L2 learners (e.g. varying proficiency levels or L1 background);
2) full-fledged vs. miniature languages;
3) natural vs. artificial language studies (where, of course, the latter are always miniature).

We believe that these approaches can be used in conjunction to test the hypothetical transitions between the various stages outlined in Table 2. In our view, what is needed in addition to the already prevalent use of between-participant designs, are more longitudinal electrophysiological studies of L2 acquisition following learners as they progress through multiple levels of L2 proficiency. Such paradigms would also help us to better understand the variables that determine the rate at which learners' progress through these stages and the end points they ultimately reach. Such variables may include properties of the language to be acquired (e.g. whether processing strategies can be transferred from the L1), of the learning environment (e.g. classroom vs. immersion based learning) and of the individual (e.g. working memory capacity).

Most previous L2 studies have focused on natural languages. However, in L2 learners of natural languages, significant increases in proficiency towards native-like levels often take months or years, are difficult to predict, and are thus difficult to study within the same participants. This is where miniature languages such as those used by Friederici *et al.* (2002), Mueller *et al.* (2005; 2007) and Morgan-Short *et al*. (2007), have significant advantages, since they allow the study of progress from novice to 'native-like' proficiency within a feasible time frame (usually a few weeks). Miniature versions of natural languages (e.g. Mueller *et al.*, 2005; 2007) have the additional advantage that real native speakers are available who can serve as controls. However, even in these approaches, it is difficult to avoid confounds of phonology and morphosyntax (as shown and discussed in Mueller *et al.*, 2005; 2007; for discussion relevant to this issue, see also Goad and White, 2004; 2006). In contrast, artificial miniature languages, if modelled after natural languages, can avoid or minimize such confounds. Some researchers have expressed doubt regarding the relevance of these paradigms to the study of natural language processing (Osterhout *et al.*, 2006). However, in our view these paradigms offer a number of important advantages that nicely complement other approaches. Specifically, the study of artificial languages allow us to control-for/eliminate problematic confounds arising in natural language studies, and even in miniature versions of them (e.g. certain problems in morphosyntax have been shown to arise as a result of phonological/prosodic problems, see Goad and White, 2004; 2006; Mueller, 2006). Furthermore, these paradigms allow the researcher to study the acquisition of specific L2 properties within the same participants from novice to native-like levels within a feasible time frame. Testing equivalent ranges of L2 proficiency levels in natural language learners, though possible in principle, is quite often far less feasible in practice. These considerations weigh strongly in favour of much further exploration and development of these 'test-tube' language learning paradigms, alongside both between-/within-group-longitudinal studies, in order to determine whether the tentative picture of possible systematic changes in ERP profiles tracking stages of L2 acquisition that we have sketched here will be empirically sustainable. Although some consider it reasonable to object to the relevance of findings based on artificial language learning paradigms as not being directly relevant to our understanding of natural language learning, it is important to remind ourselves that ERP studies using these paradigms have shown that violations comparable to those that have been tested with natural languages elicit comparable effects (e.g. LAN/P600 profile for grammatical errors in high proficiency learners of BROCANTO;

Friederici *et al.*, 2002). Faced with this evidence, one may reach one of two reasonable conclusions:

- artificial language paradigms do tap the same underlying mechanisms as those involved in natural language; or
- ERP effects like LAN and P600 are not specific to natural language.

In either case, it seems clear to us that these artificial paradigms have been empirically shown to be relevant.

In sum, although each of the various experimental approaches discussed above entail both strengths and weaknesses, when used in conjunction with one another they offer the potential to make great strides towards elucidating the neurocognitive bases of a late acquired L2.

Acknowledgements
We would like to thank Fred Genesee for valuable discussions related to the ideas presented here, and Wing-Yee Chow for assistance in data collection and analysis. This work was supported by grants awarded to the first author by the Canada Research Chair program and the Canada Foundation for Innovation (CRC/CFI; project # 201876), the Natural Sciences and Engineering Research Council of Canada (NSERC; # RGPGP 312835), and the Center for Research on Language, Mind and Brain (CRLMB; #207839).

V References

Abutalebi, J., Cappa, S.F. and **Perani, D.** 2001: The bilingual brain as revealed by functional neuroimaging. *Bilingualism* 4, 179–90.

Bahlmann, J., Gunter, T.C. and **Friederici, A.D.** 2006: Hierarchical and linear sequence processing: an electrophysiological exploration of two different grammar types. *Journal of Cognitive Neuroscience* 18, 1829–42.

Baldwin, K.B. and **Kutas, M.** 1997: An ERP analysis of implicit structured sequence learning. *Psychophysiology* 34, 74–86.

Birdsong, D., editor, 1999: *Second language acquisition and the critical period hypothesis*. Mahwah, NJ: Lawrence Erlbaum.

Birdsong, D. and **Molis, M.** 2001: On the evidence for maturational constraints in second-language acquisition. *Journal of Memory and Language* 44, 235–49.

Bley-Vroman, R. 1989: What is the logical problem of foreign language learning? In Gass, S.M. and Schacter, J., editors, *Linguistic perspectives* *on second language acquisition*. Cambridge: Cambridge University Press, 41–68.

Bowden, H.W., Sanz, C.S., Steinhauer, K. and **Ullman, M.T.** 2007: An ERP study of proficiency in second language. *Journal of Cognitive Neuroscience*, supplement, 170.

Clahsen, H. and **Felser, C.** 2006a: Grammatical Processing in Language Learners. *Applied Psycholinguistics* 27, 3–42.

—— 2006b: How native-like is non-native language processing? *Trends in Cognitive Sciences* 10, 564–70.

Coulson, S., King, J.W. and **Kutas, M.** 1998: Expect the unexpected: event-related brain response to morphosyntactic violations. *Language and Cognitive Processes* 13: 21–58.

Dehaene, S., Dupoux, E., Mehler, J., Cohen, L., Paulesu, E., Perani, D., van de Moortele, P.F., Lehericy, S. and **LeBihan, D.** 1997: Anatomical variability in the cortical representation of first and second language. *NeuroReport* 8, 3809–15.

DeKeyser, R.M. 2003: Implicit and explicit learning. In Doughty, C. and Long, M.H., editors, *The handbook of second language acquisition*. Malden, MA: Blackwell, 313–47.

DeKeyser, R.M. 2005: What makes learning second-language grammar difficult? A review of issues. *Language Learning* 55, supplement 1, 1–25.

DeKeyser, R.M. and **Larson-Hall, J.** 2002: What does the critical period really mean? In Kroll, J.F. and De Groot, A.M.B., editors, *Handbook of bilingualism: psycholinguistic approaches*. Oxford: Oxford University Press.

Doughty, C.J. and **Long, M.H.** 2003: *The handbook of second language acquisition*. Oxford: Blackwell

Eubank, L. and **Gregg, K.R.** 1999: Critical periods and (second) language acquisition: divide et impera. In Birdsong, D., editor, *Second language acquisition and the critical period hypothesis*. Hillsdale, NJ: Erlbaum, 65–99.

Flege, J.E. 1999: Age of learning and second language speech. In Birdsong, D., editor, *Second language acquisition and the critical period hypothesis*. Mahwah, NJ: Lawrence Erlbaum, 101–31.

Flege, J.E., Yeni-Komshian, G.H. and **Liu, S.** 1999: Age constraints on second-language acquisition. *Journal of Memory and Language* 41, 78–104.

Friederici, A.D. 2002: Towards a neural basis of auditory sentence processing. *Trends in Cognitive Sciences* 6, 78–84.

Friederici, A.D., Pfeifer, E. and **Hahne, A.** 1993: Event-related brain potentials during natural speech processing: effects of semantic, morphological and syntactic violations. *Cognitive Brain Research* 1, 183–92.

Friederici, A.D., Steinhauer, K. and **Frisch, S.** 1999: Lexical integration: sequential effects of syntactic and semantic information. *Memory and Cognition* 27, 438–53.

Friederici, A.D., Steinhauer, K. and **Pfeifer, E.** 2002: Brain signatures of artificial language processing: evidence challenging the critical period hypothesis. *Proceedings of the National Academy of Sciences* 99, 529–34.

Friederici, A.D., Bahlmann, J., Heim, S., Schubotz, R.I. and **Anwander, A.** 2006: The brain differentiates human and non-human grammars: functional localization and structural connectivity. *Proceedings of the National Academy of Sciences of the United States of America* 103(7), 2458–63.

Goad, H. and **White, L.** 2004: Ultimate attainment of L2 inflection: effects of L1 prosodic structure. In Foster-Cohen, S., Sharwood Smith, M., Sorace, A. and Ota, M., editors, *EUROSLA Yearbook 4*. Amsterdam: John Benjamins, 119–45.

—— 2006: Ultimate attainment in interlanguage grammars: a prosodic approach. *Second Language Research* 22, 243–68.

Gunter, T.C., Friederici, A.D. and **Schriefers, H.** 2000: Syntactic gender and semantic expectancy: ERPs reveal early autonomy and late interaction. *Journal of Cognitive Neuroscience* 12, 556–68.

Hagoort, P., Wassenaar, M. and **Brown, C.** 2003a: Syntax-related ERP-effects in Dutch. *Cognitive Brain Research* 16, 38–50.

—— 2003b: Real-time semantic compensation in patients with agrammatic comprehension: Electrophysiological evidence for multiple-route plasticity. *Proceedings of the National Academy of Sciences* 100, 4340–45.

Hahne, A. 2001: What's different in second-language processing? Evidence from event-related brain potentials. *Journal of Psycholinguistics research* 30, 251–66.

Hahne, A. and **Friederici, A.D.** 1999: Electrophysiological evidence for two steps in syntactic analysis: early automatic and late controlled processes. *Journal of Cognitive Neuroscience* 11, 194–205.

Hahne, A. and **Jescheniak, J.** 2001: What's left if the Jabberwock gets the semantics? An ERP investigation into semantic and syntactic processes during auditory sentence comprehension. *Cognitive Brain Research* 11, 199–212.

Hahne, A., Mueller, J.L. and **Clahsen, H.** 2006: Morphological processing in a second language: behavioral and event-related brain potential evidence for storage and decomposition. *Journal of Cognitive Neuroscience* 18, 121–34.

Hakuta, K., Bialystok, E. and **Wiley, E.** 2003: Critical evidence: a test of the critical-period hypothesis for second-language acquisition. *Psychological Science* 14, 31–38.

Halgren, E., Dhond, R., Christensen, N., Van Petten, C., Marinkovic, K., Lewine, J.D. and **Dale, A.M.** 2002: N400-like MEG responses modulated by semantic context, word frequency, and lexical class in sentences. *NeuroImage* 17, 1101–16.

Harley, B. and **Wang, W.** 1997: The critical period hypothesis: where are we now? In de Groot, A.M.B. and Kroll, J.F., editors, *Tutorials in bilingualism Psycholinguistic Perspectives*. Mahwah, NJ: Erlbaum, 19–51.

Hyltenstam, K. and **Abrahamsson, H.** 2003: Maturational constraints in SLA. In Doughty, C. and Long, M.H., editors, *The handbook of second language acquisition*. Malden, MA: Blackwell, 539–88.

Johnson, J.S. and **Newport, E.L.** 1989: Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology* 21, 60–99.

Kaan, E., Harris, A. Gibson, E. and **Holcomb, P.** 2000: The P600 as an index of syntactic integration difficulty. *Language and Cognitive Processes* 15, 159–201.

Knoesche, T.R., Maess, B. and **Friederici, A.D.** 2000: Processing of syntactic information monitored by brain surface current density mapping based on MEG. *Brain Topography* 12, 75–87.

Krashen, S.D., Long, M. and **Scarcella, R.** 1979: Age, rate, and eventual attainment in second language acquisition. *TESOL Quarterly* 13, 573–82.

Kubota, M., Ferrari, P. and **Roberts, T.P.L.** 2004: Human neuronal encoding of English syntactic violations as revealed by both L1 and L2 speakers. *Neuroscience Letters* 368, 235–40.

Kutas, M. and **Hillyard, S.A.** 1980: Reading senseless sentences: brain potentials reflect semantic incongruity. *Science* 207, 203–05.

Long, M.H. 1990: Maturational constraints on language development. *Studies in second language acquisition* 12, 251–85.

Marchman, V.A. 1993: Constraints on plasticity in a connectionist model of the English past tense. *Journal of Cognitive Neuroscience* 5, 215–34.

Martin-Loeches, M., Munoz, F., Casado, P., *et al.* 2005: Are the anterior negativities to grammatical violations indexing working memory? *Psychophysiology* 42, 508–19.

McLaughlin, J., Osterhout, L. and **Kim, A.** 2004: Neural correlates of second-language word learning: minimal instruction produces rapid change. *Nature Neuroscience* 7, 703–704.

Morgan-Short, K., Steinhauer, K., Sanz, C. and **Ullman, M.T.** 2007: An ERP investigation of second language processing: effects of proficiency and explicit and implicit training. *Journal of Cognitive Neuroscience*, supplement, 164.

Mueller, J.L. 2006: L2 in a nutshell: the investigation of second language processing in the miniature language model. *Language Learning* 56, 235–70.

Mueller, J.L., Hahne, A., Fujii, Y. and **Friederici, A.D.** 2005: Native and nonnative speakers' processing of a miniature version of Japanese as revealed by ERPs. *Journal of Cognitive Neuroscience* 17, 1229–44.

Mueller, J.L., Hirotani, M. and **Friederici, A.D.** 2007: ERP evidence for different strategies in the processing of case markers in native speakers and non-native learners. *BMC Neuroscience* 8(18).

Neville, H., Nicol, J.L., Barss, A., Forster, K.I. and **Garrett, M.F.** 1991: Syntactically based sentence processing classes: evidence from event-related brain potentials. *Journal of Cognitive Neuroscience* 3, 151–65.

Newport, E.L. 1993: Maturational constraints on language learning. In P. Bloom, editor, *Language Acquisition*. Cambridge, MA: MIT Press, 543–60.

Ojima, S., Nakata, H. and **Kakigi, R.** 2005: An ERP study on second language learning after childhood: effects of proficiency. *Journal of Cognitive Neuroscience* 17, 1212–28.

Osterhout, L. and **Holcomb, P.J.** 1992: Event-related brain potentials elicited by syntactic anomaly. *Journal of Memory and Language* 31, 785–806.

Osterhout, L., McLaughlin, J., Pitkanen, I., Frenck-Mestre, C. and **Molinaro, N.** 2006: Novice learners, longitudinal designs, and event-related potentials: a means for exploring the neurocognition of second language processing. *Language Learning* 56, 199–230.

Osterhout, L. and **Mobley, L.A.** 1995: Event-related brain potentials elicited by failure to agree. *Journal of Memory and Language* 34, 739–73.

Pakulak, E. and **Neville, H.** 2004: Individual differences in online syntactic processing in monolingual adults as reflected by ERPs. In *Proceedings of the 17th Annual CUNY Conference on Human Sentence Processing*, 149. Available online at: http://www.ling.umd.edu/cuny2004/CUNY_Handbook_2004.pdf (October 2008).

Perani, D., Paulesu, E., Galles, N.S., Dupoux, E., Dehaene, S., Bettinardi, V., Cappa, S.F., Fazio, F. and **Mehler, J.** 1998: The bilingual brain: proficiency and age of acquisition of the second language. *Brain* 12, 1841–52.

Phillips, C., Kazanina, N. and **Abada, S.H.** 2005: ERP effects of the processing of syntactic long-distance dependencies. *Cognitive Brain Research* 22, 407.

Pinker, S. and **Prince, A.** 1994: Regular and irregular morphology and the psychological status of rules of grammar. In Lima, S.D., Corrigan, R.L. and Iverson, G.K., editors, *The reality of linguistic rules*. Philadelphia, PA: John Benjamins.

Reber, A.S. 1967: Implicit learning of artificial grammars. *Journal of Verbal Learning and Verbal Behavior* 6, 855–63.

Reber, P.J. and **Squire, L.R.** 1999: Intact learning of artificial grammars and intact category learning by patients with Parkinson's disease. *Behavioral Neuroscience* 113, 235–42.

Rossi, S., Gugler, M.F., Friederici, A.D. and **Hahne, A.** 2006: The impact of proficiency on syntactic second-language processing of German and

Italian: evidence from event-related potentials. *Journal of Cognitive Neuroscience* 18, 2030–48.

Steinhauer, K. and **Connolly, J.F.** 2008: Event-related potentials in the study of language. In Stemmer, B. and Whitaker, H.A., editors, *Handbook of the neuroscience of language.* New York: Elsevier, 91–104.

Steinhauer, K., White, E., Cornell, S., Genesee, F. and **White, L.** 2006: The neural dynamics of second language acquisition: evidence from Event-Related Potentials. *Journal of Cognitive Neuroscience*, supplement, 99.

Ullman, M.T. 2001: The neural basis of lexicon and grammar in first and second language: the declarative/procedural model. *Bilingualism: Language and Cognition* 4, 105–22.

—— 2005: A cognitive neuroscience perspective on second language acquisition: the declarative/procedural model. In Sanz, C., editor, *Mind and context in adult second language acquisition: methods, theory and practice*. Washington, DC: Georgetown University Press, 141–78.

Van Petten, C. and **Luka, B.J.** 2006: Neural localization of semantic context effects in electromagnetic and hemodynamic studies. *Brain and Language* 97, 279–93.

Wartenburger, I., Heekeren, H.R., Abutalebi, J., Cappa, S.F., Villringer, A. and **Perani, D.** 2003: Early setting of grammatical processing in the bilingual brain. *Neuron* 37, 159–70.

Weber-Fox, C.M. and **Neville, H.J.** 1996: Maturational constraints on functional specializations for language processing: ERP and behavioral evidence in bilingual speakers. *Journal of Cognitive Neuroscience* 8, 231–56.

Weber-Fox, C., Davis, L.J. and **Cuadrado, E.** 2003: Event-related brain potential markers of high-language proficiency in adults. *Brain and Language* 85, 231–44.

White, E.J., Genesee, F., Drury, J.E. and **Steinhauer, K.** 2007: Before and after: an ERP investigation of late second language learning in an intensive language course. *Journal of Cognitive Neuroscience*, supplement, 290.

White, L. 2003: On the nature of the interlanguage representation: universal grammar in the second language. In Doughty, C. and Long, M.H., editors, *The handbook of second language acquisition*. Malden, MA: Blackwell, 19–42.

White, L. and **Genesee, F.** 1996: How native is near-native? The issue of ultimate attainment in adult second language acquisition. *Second Language Research* 12, 233–65.