

# Second Language Research

<http://slr.sagepub.com>

---

## **Event-related brain potentials and second language learning: syntactic processing in late L2 learners at different L2 proficiency levels**

Janet G. van Hell and Natasha Tokowicz  
*Second Language Research* 2010; 26; 43  
DOI: 10.1177/0267658309337637

The online version of this article can be found at:  
<http://slr.sagepub.com/cgi/content/abstract/26/1/43>

---

Published by:



<http://www.sagepublications.com>

**Additional services and information for *Second Language Research* can be found at:**

**Email Alerts:** <http://slr.sagepub.com/cgi/alerts>

**Subscriptions:** <http://slr.sagepub.com/subscriptions>

**Reprints:** <http://www.sagepub.com/journalsReprints.nav>

**Permissions:** <http://www.sagepub.co.uk/journalsPermissions.nav>

**Citations** <http://slr.sagepub.com/cgi/content/refs/26/1/43>

# Event-related brain potentials and second language learning: syntactic processing in late L2 learners at different L2 proficiency levels

**Janet G. van Hell** *Radboud University Nijmegen, The Netherlands and The Pennsylvania State University, USA* and

**Natasha Tokowicz** *University of Pittsburgh, USA*

Received April 2008; revised July 2008; accepted September 2008

There are several major questions in the literature on late second language (L2) learning and processing. Some of these questions include: Can late L2 learners process an L2 in a native-like way? What is the nature of the differences in L2 processing among L2 learners at different levels of L2 proficiency? In this article, we review studies that addressed these questions using event-related brain potentials (ERPs) in late learners and that focused on syntactic processing. ERPs provide an on-line, millisecond-by-millisecond record of the brain's electrical activity during cognitive processing. ERP measures can thus provide valuable information on the timing and degree of neural activation as language processing (here: syntactic processing in L2) unfolds over time. After discussing the use of ERPs for the study of L2 learning and processing, we review electrophysiological studies on syntactic and morphosyntactic processing in late L2 learners with different levels of L2 proficiency. The currently available evidence indicates that patterns of neural activity in the brain during syntactic and morphosyntactic processing can be modulated by various, possibly inter-related, factors including the similarity or dissimilarity of syntactic structures in L2 and L1, the exact nature of the syntactic structure L2 learners seek to comprehend and the concomitant expectancies

Address for correspondence: Janet G. van Hell, Behavioural Science Institute, PO Box 9104, Radboud University Nijmegen, 6500 HE Nijmegen, The Netherlands; email: J.vanHell@pwo.ru.nl

© The Author(s), 2010. Reprints and permissions:

<http://www.sagepub.co.uk/journalspermissions.nav>

10.1177/0267658309337637

they can generate with regard to violations in this structure, and the L2 learners' level of L2 proficiency. Together these studies show that ERPs can successfully elucidate subtle differences in syntactic processing between L2 learners and native speakers, and among L2 learners at different levels of L2 proficiency, which are difficult to detect or that might have remained undetected with behavioural measures.

**Keywords:** event-related brain potentials, late L2 learners, syntactic processing in late L2 learners, age effect in second language learning, electroencephalogram, violation paradigm, grammaticality judgements, critical period hypothesis, competition model for L2 learning, L2 processing, automaticity in L2 processing, integration and repair processes in L2 processing, impact of level of L2 proficiency on language processing

## **I Introduction**

One of the 125 critical questions for the next 25 years, as compiled by *Science* for its 125th anniversary (Kennedy and Norman, 2005), pertains to the biological basis of second language (L2) learning as revealed by the monitoring of brain activity. This research question ensued from the observation that 'children pick up languages with ease while adults often struggle to learn train-station basics in a foreign language' (Kennedy and Norman, 2005: 93). Of course, this observation is not new: it has fascinated generations of researchers from linguistics, psychology and education who share an interest in L2 learning. In the past decade, researchers have begun to address this topic from a neurocognitive perspective. In this review article, we discuss a selected set of articles that used one particular neurocognitive method, event-related brain potentials (ERPs), to study L2 learning (for reviews on other neuroimaging techniques and L2 learning, see Indefrey, 2006; Abutalebi and Green, 2007). We specifically focus on syntactic processing in late L2 learners, whom we define as individuals who learned their L2 in middle childhood (around 8–10 years) or later, well after adequate L1 language skills had been achieved.

After we discuss the basic principles of the ERP technique, we review ERP studies on (morpho)syntactic processing in late L2 learners with different levels of L2 proficiency (for an overview of ERP studies on bilingual phonology, semantics and language control, see Moreno *et al.*, 2008; for an overview of ERP studies on language switching, see

van Hell and Witteman, 2009). The studies we review address several interrelated questions:

- Can late L2 learners process an L2 in a native-like way?
- If not, what is the nature of the differences between native speakers and late L2 learners?
- How do these differences change with proficiency?
- What factors influence L2 learners' syntactic and morphosyntactic processing in L2?

## II Basic principles of ERPs and language research

The variation of electrical activity produced by large populations of brain cells can be measured by electrodes placed on the scalp. The record of voltage variations over time is called the electroencephalogram (EEG). ERPs are derived from the EEG and represent voltage changes in electrical brain activity that are time-locked to an external event like the presentation of a word (for excellent introductions to ERP recordings and analyses, see Fabiani *et al.*, 2000; Handy, 2005; Luck, 2005). ERPs provide a record of the brain's electrical activity during mental processing as it unfolds over time. ERPs can thus be used to index perceptual and cognitive processes involved in language comprehension and production.

The typical ERP signal consists of a series of positive and negative peaks, termed components, related to stimulus processing. ERP components are characterized by polarity, latency, amplitude, topographic distribution across the scalp, and a functional description of the experimental effects they index. A component either has a positive polarity (positive-going wave, labelled by P) or negative polarity (negative-going wave, labelled by N). Latency measures reflect the time course of the signal, and include onset latency (when a component begins) and peak latency (when a component reaches its peak amplitude). Components are often labelled in terms of their polarity and the latency at which their amplitude reaches its maximum (e.g. P600, as will be explained below). A component's relative amplitude is assumed to reflect the degree of engagement of the associated cognitive processes. Topographic distribution pertains to the typical scalp distribution of the ERP component. Two components similar in polarity and latency that differ in terms of scalp distribution are taken to reflect different processes. Finally, components are functionally described in terms of the experimental manipulation

to which a component is sensitive and the cognitive process(es) the component is assumed to reflect.

The main components considered to index particular aspects of adult native-language processing are the early left anterior negativity (ELAN), the left anterior negativity (LAN), the N400 and the P600. The ELAN, which is an anterior negativity that occurs in the 150–250 ms latency range, is often lateralized over the left hemisphere, and is assumed to reflect syntactic-structure building that occurs extremely rapidly (e.g. Hahne and Friederici, 1999). The LAN is also an anterior negativity that is often left-lateralized, but it occurs slightly later, in the 300–500 ms range, indexes syntactic processing and appears to correlate particularly with morphosyntactic errors (e.g. Friederici, 2002). The classic N400 is a large-amplitude, negative-going wave beginning about 300 ms post-stimulus and reaching its maximum around 400 ms post-stimulus. Although the N400 occurs in the same time window as the LAN, the N400 is usually largest at central and parietal electrode sites and indexes the integration of meaning and world knowledge (e.g. Kutas and Hillyard, 1980; Hagoort *et al.*, 2004); its amplitude increases as the strength of the semantic relation between a target and the preceding sentence decreases. The P600 is a positive-going wave appearing around 500–600 ms post-stimulus and extending for several hundred milliseconds (e.g. Osterhout and Holcomb, 1992; Hagoort *et al.*, 1993). It has a broad posterior scalp distribution and is greatest over centroparietal regions. The P600 is sensitive to (morpho) syntactic processing and occurs for syntactic violations (e.g. Osterhout and Holcomb, 1992) or syntactically complex structures (Kaan *et al.*, 2000).

These different ERP components have been attributed to functionally different stages in language processing (e.g. Hahne and Friederici, 1999; Friederici, 2002): the ELAN indexes an initial, first-pass automatic structure-building process (i.e. identification of word category), the LAN/N400 indexes the integration of semantic and morphosyntactic information with the goal of thematic role assignment, and the P600 indexes a late controlled process of re-analysis and repair of effortful syntactic integration. So, the ELAN is assumed to reflect highly automatic, first-pass parsing, and the P600 is assumed to reflect more controlled, second-pass parsing.

Note that these components' characteristics of polarity, latency, peak amplitude and topographic distribution are based on research with adult native-language speakers. The increasing body of research on less proficient speakers of a language, including L2 learners and children,

indicates that the ‘typical’ characteristics of components may be slightly different in these groups. For example, Hahne (2001) observed a delayed P600 in L2 (we elaborate on this at a later point in this article).

The majority of studies on (morpho)syntactic processing have used the violation paradigm. In many of these studies, ERPs are time-locked to the presentation of the ‘critical’ or violating word. For example, in the sentence \**The man walk on the beach*, the critical word is ‘walk’, because this is the earliest point at which participants could detect the violation. Across participants, performance on this sentence would be compared to performance to the same word in a sentence without the violation. It is generally preferable to compare activity to the same exact stimulus word in two different contexts, although this is not always possible. When processing in response to different stimuli is compared, it is important to attempt to match those stimuli as closely as possible on other word-specific characteristics (e.g. frequency, length) that may affect the responses (see Tokowicz and Warren, 2008).

What is the advantage of using ERPs to study L2 learning? ERPs provide valuable information on the timing and degree of neural activation as language processing unfolds over time and provide insight into the physiological correlates of language behaviour. ERPs are particularly useful in the study of L2 learning and processing because they do not require that overt responses be made. They have also been shown to indicate learning even when overt, behavioural measures have not (e.g. McLaughlin *et al.*, 2004; Tokowicz and MacWhinney, 2005).

Furthermore, ERPs can indicate whether there are qualitative and/or quantitative differences between L2 learners and native speakers, or for various types of L2 violations. In particular, qualitative differences are indicated by the presence/absence of some ERP component (e.g. finding an N400 rather than a P600). By contrast, quantitative differences are measured in terms of the absolute size of a particular effect (e.g. whether learners show a P600 of the same magnitude as native speakers) or in terms of the timing of a component (e.g. whether learners have a delayed P600 onset or peak latency). See Handy (2005) for a discussion of additional advantages and some disadvantages of the ERP approach.

In the remainder of this article, we review ERP studies that address syntactic processing in late L2 learners and contribute to answering the leading questions in this article. Before reviewing the ERP studies, we briefly describe several perspectives that formed the theoretical background for these studies.

### III Can late L2 learners process an L2 in a native-like way?

#### 1 *Theoretical perspectives*

One view on late L2 learning describes a critical period for L2 acquisition (see, for example, Weber-Fox and Neville, 1996). According to this hypothesis, the critical period in normal language acquisition depends on diminished brain capacity, which affects the ultimate success of learning an L2 (for an extensive overview, see, for example, Lenneberg, 1967; see Birdsong, 1999). The critical period hypothesis suggests that late learners of an L2 cannot attain native-like proficiency; L2 phonology and syntax are thought to be more vulnerable to age of exposure than vocabulary. On this view, there is an inflection in the relationship between age of acquisition and ultimate attainment (e.g. a strong negative linear relationship between age of acquisition and ultimate attainment during the critical period, followed by a weaker relationship between age of acquisition and attainment outside the period; see Birdsong, 1999).

An alternative to the critical period hypothesis is the assumption of an age-related decline in L2 acquisition that is not due to a particular period of brain development. Instead, this view emphasizes a linear decline in ultimate attainment with increased age of first exposure that is attributable to decline in general cognitive mechanisms (e.g. Birdsong, 1999). On this view, there is no endpoint to the age-related decline in attainment, although there may be some (pre- or post-maturational) beginning point for the decline.

The Competition Model for L2 learning (e.g. MacWhinney, 1997) provides a specific explanation for an age-related decline in L2 acquisition related to the entrenchment of the L1 system. This model emphasizes the strength of cues used in language processing; the strength is based partly on the reliability and availability of cues in the input to the learner. The more exposure to a language, the stronger the relevant cues will become. Stronger cues will be used earlier in learning (e.g. McDonald, 1987). Therefore, when L1 is entrenched, L2 learning will be more difficult, except for aspects that are similar in the two languages. This is because when the two languages are similar, there will be no competition between cues in L1 and L2, which enables L2 learners to transfer knowledge of their L1 to L2. However, when the two languages are dissimilar, they compete, which may result in negative transfer. And, for some syntactic structures that are unique in the learner's L2, there is no

competition from L1, in which case the availability and validity of cues in L2 will determine how well they will be learned.

Other views describe the development of L2 proficiency in terms of automaticity in L2 processing (e.g. Ullman, 2001). Ullman's declarative/procedural model proposes that L1 users employ a declarative system to process lexical information and a procedural system to process rule-based syntactic information. In extension to L2 learning, the procedural system is thought to have very limited applicability, particularly in low proficiency or beginning learners. On this view, increased L2 proficiency leads to a qualitative shift in the type of processing that is done, and in the brain systems that subserve processing.

These various views make predictions about factors that affect ultimate L2 attainment. They vary in the degree to which they provide alternatives for what will change with increased L2 proficiency. For example, the critical period hypothesis predicts qualitative differences between learners who are first exposed during vs. after the critical period. The Competition Model and age-related decline framework predict that later exposure will lead to slower learning, added competition or lower ultimate attainment. However, they do not specifically predict qualitative differences as a function of age of exposure, although beginning learners may demonstrate insensitivity to violations in syntactic structures that differ across languages. The declarative/procedural model predicts a qualitative shift in processing to more implicit/procedural systems for syntactic processing if an individual becomes highly proficient.

## 2 Review of ERP studies

We now review ERP studies that address whether late L2 learners can process L2 in a native-like way. We divide this discussion into two parts. We first discuss studies that focus on grammatical violations at the sentence level, in particular violations of phrase structure. We then discuss studies that examined violations of morphology, in particular violations of participial inflection, noun plurals, subject–verb agreement, determiner number and gender agreement, and auxiliary omission. Table 1 presents an overview of the studies we discuss, listing the linguistic structure, the type of learners that were examined and the main findings.

*a Grammatical violations of phrase structure:* The first ERP study that focused on syntactic processing in L2 learners was Weber-Fox and Neville's (1996) study of Chinese–English bilinguals who were

**Table 1** Overview of the ERP studies on (morpho) syntactic violations in late L2 learners available in the literature and reviewed in this article

Study	Linguistic structure (presentation mode)	Participant group (proficiency level and self-rated proficiency)	Main findings
<i>Grammatical violations of phrase structure:</i>			
Weber-Fox and Neville (1996)	phrase structure (visual)	<ul style="list-style-type: none"> <li>native speakers of English</li> <li>Chinese learners of L2 English</li> <li>first exposed to L2 at age 1–3, 4–6, 7–10, 11–13, or &gt; 16</li> <li>self-ratings: 3.8, 3.8, 3.7, 3, 1, 2.6 out of 4</li> </ul>	<ul style="list-style-type: none"> <li>early (N125) and later left-lateralized negativity (between 300–500 ms), followed by P600</li> <li>no early left-lateralized negativity (N125)</li> <li>later negativity (between 300–500 ms), bilateral in learners exposed to L2 &gt; 11 years</li> <li>P600 in learners exposed to L2 &lt; age 10</li> <li>delayed positivity beginning at 700 ms in learners exposed to L2 11–13 years</li> <li>no P600 in learners exposed to L2 &gt; 16</li> </ul>
	specificity constraint (visual)	<ul style="list-style-type: none"> <li>native speakers of English</li> </ul>	<ul style="list-style-type: none"> <li>early (N125) and sustained left-lateralized negativity between 300–500 ms and 500–700 ms</li> </ul>
		<ul style="list-style-type: none"> <li>Chinese learners of L2 English</li> <li>first exposed to L2 at age 1–3, 4–6, 7–10, 11–13 or &gt; 16</li> <li>self-ratings: 3.8, 3.8, 3.7, 3.1, 2.6 out of 4</li> </ul>	<ul style="list-style-type: none"> <li>early (N125) and sustained left-lateralized negativity between 300–500 ms and 500–700 ms in learners exposed to L2 before age 11</li> <li>bilateral early negativity (N125) and weak sustained left-lateralized negativity between 300–500 ms and 500–700 ms in learners exposed to L2 between 11–13</li> <li>no early or sustained negativity in learners exposed to L2 &gt; 16 years.</li> </ul>

Hahne and Friederici (2001)	Phrase structure (auditory)	<ul style="list-style-type: none"> <li>• native speakers of German</li> <li>• Japanese late learners of German</li> <li>• moderately proficient</li> <li>• self-ratings: 3.5 out of 6</li> <li>• native speakers of German</li> </ul>	<ul style="list-style-type: none"> <li>• early anterior negativity followed by P600</li> <li>• no marked sensitivity to syntactic violations</li> </ul>
Hahne (2001)	Phrase structure (auditory)	<ul style="list-style-type: none"> <li>• Russian late learners of German</li> <li>• proficient</li> <li>• self-ratings: 3.4 out of 4</li> <li>• native speakers of French</li> </ul>	<ul style="list-style-type: none"> <li>• early anterior negativity followed by P600</li> <li>• no early anterior negativity</li> <li>• P600 with delayed peak onset</li> </ul>
Isel (2007)	Phrase structure (auditory)	<ul style="list-style-type: none"> <li>• native speakers of French</li> </ul>	<ul style="list-style-type: none"> <li>• ELAN between 150–300 ms followed by frontal negativity between 300–600 ms and a P600</li> </ul>
Rossi <i>et al.</i> (2006)	Phrase structure (auditory)	<ul style="list-style-type: none"> <li>• German late learners of French</li> <li>• proficient</li> <li>• self-ratings: not provided</li> <li>• native speakers of German and Italian</li> <li>• Italian learners of German and German learners of Italian</li> <li>• both proficient</li> <li>• self-ratings: 4.5 out of 6</li> </ul>	<ul style="list-style-type: none"> <li>• ELAN between 150–300 ms followed by frontal negativity between 300–600 ms, and bilaterally distributed late anterior negativity between 750–1500 ms</li> <li>• ELAN, additional negativity between 250 and 650 ms, P600</li> <li>• ELAN, additional negativity between 250 and 650 ms, P600</li> </ul>

(Continued)

Table 1 (Continued)

Study	Linguistic structure (presentation mode)	Participant group (proficiency level and self-rated proficiency)	Main findings
<i>Violations of morphosyntaxis:</i>			
Rossi <i>et al.</i> (2006)	subject-verb agreement (auditory)	<ul style="list-style-type: none"> <li>Italian learners of German and German learners of Italian</li> <li>both moderately proficient</li> <li>self-ratings: 3.0 out of 6</li> </ul>	<ul style="list-style-type: none"> <li>ELAN, additional negativity between 250 and 600 ms, and a P600 with a delayed onset and reduced amplitude</li> </ul>
		<ul style="list-style-type: none"> <li>native speakers of German and Italian</li> </ul>	<ul style="list-style-type: none"> <li>LAN, followed by P600</li> </ul>
		<ul style="list-style-type: none"> <li>Italian learners of German</li> <li>German learners of Italian</li> <li>both proficient</li> <li>self-ratings: 4.5 out of 6</li> </ul>	<ul style="list-style-type: none"> <li>LAN, followed by P600</li> </ul>
		<ul style="list-style-type: none"> <li>Italian learners of German, German learners of Italian, both moderately proficient</li> <li>self-ratings: 3.0 out of 6</li> </ul>	<ul style="list-style-type: none"> <li>No LAN, only P600 (with delayed onset and reduced amplitude compared to proficient speakers)</li> </ul>
Chen <i>et al.</i> (2007)	subject-verb agreement (visual)	<ul style="list-style-type: none"> <li>native speakers of English</li> </ul>	<ul style="list-style-type: none"> <li>LAN, followed by P600</li> </ul>
		<ul style="list-style-type: none"> <li>Chinese learners of English</li> <li>moderately proficient</li> <li>self-ratings: 4.6 out of 7</li> </ul>	<ul style="list-style-type: none"> <li>late anterior-central negativity between 500–700 ms (labelled N600 no LAN, no P600)</li> </ul>
		<ul style="list-style-type: none"> <li>native speakers of English</li> </ul>	<ul style="list-style-type: none"> <li>LAN, followed by P600</li> </ul>
Ojima <i>et al.</i> (2005)	subject-verb agreement (visual)	<ul style="list-style-type: none"> <li>Japanese learners of English</li> <li>proficient</li> <li>self-ratings: 7.4 out of 10</li> </ul>	<ul style="list-style-type: none"> <li>LAN, no P600</li> </ul>

<p>Hahne <i>et al.</i> (2006)</p>	<p>participial inflection (visual)</p>	<ul style="list-style-type: none"> <li>• Japanese learners of English</li> <li>• moderately proficient</li> <li>• self-ratings: 4.4 out of 10</li> <li>• native speakers of German</li> </ul>	<ul style="list-style-type: none"> <li>• no LAN, no P600</li> <li>• regularizations: left anterior negativity, no P600</li> <li>• irregularizations: N400</li> <li>• regularizations: bilateral anterior negativity between 250–600 ms, followed by P600 between 600–1000 ms</li> <li>• irregularizations: centrally distributed negativity between 450–600 ms ('morphosyntactic negativity')</li> </ul>
<p>Tokowicz and MacWhinney (2005)</p>	<p>noun plurals (visual)</p>	<ul style="list-style-type: none"> <li>• Russian late learners of German</li> <li>• proficient</li> <li>• self-ratings: 5 out of 6</li> <li>• native speakers of German</li> <li>• Russian late learners of German</li> <li>• proficient</li> <li>• self-ratings: 5 out of 6</li> </ul>	<ul style="list-style-type: none"> <li>• regularizations: left anterior negativity, followed by a P600</li> <li>• irregularizations: N400</li> <li>• regularizations: no early anterior negativity, but a P600</li> <li>• irregularizations: N400 between 300–800 ms</li> <li>• P600</li> </ul>
<p>Tokowicz and MacWhinney (2005)</p>	<p>auxiliary omission (visual)</p>	<ul style="list-style-type: none"> <li>• English learners of Spanish</li> <li>• low proficient</li> <li>• self-ratings: 5.3 out of 10</li> </ul>	

(Continued)

Table 1 (Continued)

Study	Linguistic structure (presentation mode)	Participant group (proficiency level and self-rated proficiency)	Main findings
Osterhout <i>et al.</i> (2006)	grammatical gender agreement (visual)	<ul style="list-style-type: none"> <li>• same</li> </ul>	<ul style="list-style-type: none"> <li>• P600</li> </ul>
	determiner number agreement (visual)	<ul style="list-style-type: none"> <li>• same</li> </ul>	<ul style="list-style-type: none"> <li>• no P600</li> </ul>
	subject–verb agreement (visual)	<ul style="list-style-type: none"> <li>• native speakers of French</li> </ul>	<ul style="list-style-type: none"> <li>• P600</li> </ul>
		<ul style="list-style-type: none"> <li>• English learners of French</li> </ul>	<ul style="list-style-type: none"> <li>• N400 after 1 month of instruction</li> </ul>
		<ul style="list-style-type: none"> <li>• low proficient</li> </ul>	<ul style="list-style-type: none"> <li>• P600 after 4 months of instruction</li> </ul>
		<ul style="list-style-type: none"> <li>• self-ratings: not provided</li> </ul>	
	determiner number agreement (visual)	<ul style="list-style-type: none"> <li>• native speakers of French</li> </ul>	<ul style="list-style-type: none"> <li>• P600</li> </ul>
		<ul style="list-style-type: none"> <li>• English learners of French</li> </ul>	<ul style="list-style-type: none"> <li>• no effect after 1, 4 or 8 months of instruction</li> </ul>
		<ul style="list-style-type: none"> <li>• low proficient</li> </ul>	
		<ul style="list-style-type: none"> <li>• self-ratings: not provided</li> </ul>	

exposed to L2 English at different ages. This study's theoretical starting point was to explore critical period effects in semantic and syntactic processing in L2 learners. The L2 learners were adults at the time of testing, but had first been exposed to English at different ages: 1–3, 4–6, 7–10, 11–13 or after 16. Participants read sentences containing semantic violations and two critical types of syntactic violations: phrase structure (e.g. *\*The scientist criticized Max's of proof the theorem*) and specificity constraint (e.g. *\*What did the scientist criticize Max's proof of?*). The ERPs of all the L2 learners to semantic violations were comparable to those of the native speakers, although the L2 learners first exposed to English at age 11 or later had slightly delayed N400 peak latencies. In contrast, ERP responses of all L2 learners to syntactic violations, irrespective of age of first exposure to English, were different from those of the native speakers. In native speakers, phrase structure violations elicited an early left-lateralized negativity (labelled N125), followed by a later left-lateralized negativity between 300–500 ms, and a P600 between 500–700 ms. By contrast, none of the L2 learners showed the early left-lateralized negativity. The L2 learners did show the later negativity between 300–500 ms, although this negativity was bilaterally distributed for the late L2 learners who had been exposed to English at age 11 or later. The P600 also differed as a function of age of exposure. The earlier L2 learners (1–3, 4–6, 7–10 groups), like the native speakers, showed a P600. The learners exposed between ages 11 and 13 showed a delayed increase in positivity beginning around 700 ms whereas the latest group of learners (after 16) showed no effect.

To the second type of syntactic violation tested by Weber-Fox and Neville (1996), namely specificity constraint violations, ERP responses of the earlier L2 learners (before age 11) were comparable to those of the native speakers: there was an early (N125) and sustained left-lateralized negativity between 300–500 ms and 500–700 ms (with the exception that the L2 learners first exposed to English between 1–3 years, surprisingly, did not show an N125). In contrast, the L2 learners first exposed to English between 11–13 showed a bilateral early negativity (N125), followed by a weak later left-lateralized negativity between 300–500 ms. The L2 learners exposed to English after age 16 showed no early (N125) or late (300–500 ms) enhanced negativity. The ERPs of the 11–13 group also showed a weak P600 between 500–700 ms, but no P600 was observed in the ERPs of the L2 learners first exposed to English after age 16.

To summarize, Weber-Fox and Neville's (1996) findings indicate that ERP responses of Chinese learners of English to semantic violations were, by and large, similar to those of native speakers. A different pattern was obtained for the syntactic violations. Although ERP responses to syntactic violations of all L2 learners, irrespective of age of first exposure to English, were different from those of native speakers, these differences were most pronounced in the late L2 learners, that is, the L2 learners first exposed to English at age 11 or later. These late L2 learners' performance on a grammaticality judgement task was also significantly worse than that of native speakers, indicating relatively lower proficiency in addition to later age of exposure. It is important to note that because proficiency and age of exposure were correlated in this study, the pattern of findings could be attributable to proficiency rather than age of exposure *per se*.

Phrase structure violations were also examined by Hahne and Friederici (2001) in their study of late learners of German. They examined the comprehension of auditorily presented passive voice sentences that contained a semantic violation (e.g. *\*Der Vulkan wurde gegessen* [*\*'The volcano was eaten'*]), a syntactic violation (e.g. *\*Das Eis wurde im gegessen* [*\*'The ice-cream was in-the eaten'*]), or both. The L2 learners were adult Japanese native speakers who had been living in Germany between 2–79 months, and who had learned German in formal settings after the age of 18. Their self-ratings and relatively high error rates in grammaticality judgements of the experimental sentences suggest that these learners were only moderately proficient in German. For semantic violations, the ERPs of the L2 learners were, by and large, similar to those of native German speakers, comparable to the findings of Weber-Fox and Neville (1996). By contrast, for the sentences containing a syntactic violation, the L2 learners' performance contrasted strongly with that of the native speakers. The L2 learners showed hardly any difference between sentences with and without a syntactic violation, whereas the native speakers showed a modulation of an early anterior negativity followed by a P600. This study indicates that semantic integration processes are similar for native speakers and L2 learners, but that syntactic processing in L2 learners differs from that of native speakers, as evidenced by the absence of syntax-related ERP components in the L2 learners. Hahne and Friederici proposed that this finding could be driven by the fact that prepositions, which were critical in the phrase structure violations of their study, are difficult for native Japanese speakers to learn because prepositions do not exist in Japanese.

In a follow-up study, Hahne (2001) tested whether Hahne and Friederici's (2001) findings would hold for learners of German with a native language that did contain prepositions, Russian. Hahne tested a highly proficient group of German learners and a group of native German speakers. Participants were auditorily presented with syntactically correct and incorrect sentences in which the participle immediately followed the preposition, thus containing a phrase structure error (e.g. \**Das Geschäft wurde am geschlossen* [\* 'The shop was being on closed']). In the native speakers, these phrase structure violations elicited an early anterior negativity between 100–250 ms, followed by a P600. The L2 learners, in contrast, showed no early anterior negativity, but they did show a P600, which peaked about 150 ms later than in the native speakers. It should be noted, however, that in addition to speaking a native language that contains prepositions, Hahne's Russian learners of German were also more proficient in German than Hahne and Friederici's Japanese learners of German. So, it is not clear whether the absence of syntax-related ERP components in the relatively low proficient Japanese learners of German was driven by the absence of prepositions in Japanese, or by their relatively low L2 German proficiency.

The ERP studies discussed so far show that, in native speakers, phrase structure violations typically evoke a biphasic ERP pattern: an ELAN, assumed to reflect highly automatic syntactic phrase-building processes, and a P600 assumed to reflect syntactic re-analysis and repair. In contrast, studies on phrase structure violations in L2 learners, as discussed above, show no ELANs (Weber-Fox and Neville, 1996; Hahne, 2001; Hahne and Friederici, 2001), suggesting that no early automatic phrase-building processes occur in L2.

An alternative explanation has recently been put forward by Isel (2007), who noted that the phrase structure violations used in previous studies on L2 processing occurred in syntactically non-obligatory constituents, hence, in optional constituents that can be present or absent in full sentences. For example, in the phrase structure example sentence of Hahne and Friederici (2001) \**Das Eis wurde im gegessen*, where a noun would be expected in the prepositional adverbial expression (e.g. *im Theater* ['in the theatre']), the preposition *im Theater* is optional: both *Das Eis wurde gegessen* ('The ice-cream was eaten') and *Das Eis wurde im Theater gegessen* ('The ice-cream was eaten in the theatre') are acceptable. In contrast, in sentences like *The driver who is in the car is sleeping*, the prepositional phrase *in the car* is obligatory. Isel proposed that automatic phrase-building processes during L2 processing are more

likely to occur when the constituent is obligatory. When obligatory, these violations should evoke an ELAN, because expectancies of the sentence parser will be stronger and violations will be easier to detect in these constituents. Isel therefore tested L2 learners' ERP responses to syntactically obligatory constituents, the predicative prepositional phrase of a subject-modifying relative clause (e.g. *\*Le chauffeur qui est dans la – dort* [*\*‘The driver who is in the – is sleeping’*]). The participants were native German-speaking highly proficient late learners of L2 French.

The learners' ERP responses to the auditorily presented sentences indeed showed an early negativity between 150–300ms, followed by a frontal negativity between 300–600ms with a maximum over the anterior region, and a bilaterally distributed later anterior negativity between 750–1500ms. No significant P600 effect was obtained (although a P600 was observable in the descriptive data). A control group of native French speakers also yielded an ELAN followed by a later frontal negativity between 300–600ms and a P600 between 600–1000ms (Isel *et al.*, 2007). This study thus indicates that the involvement of automatic structure-building processes in L2 processing may depend on the syntactic status of the constituent (i.e. obligatory or optional) in which the phrase structure (word category) violation occurs. If word category violations occur in syntactically obligatory constituents, and the parser's expectancies are high because of the obligatory status of this constituent, L2 learners' ERP responses reflect early, automatic structure building processes in L2, just as in native speakers.

The idea that the ELAN in L2 learners' response to phrase structure violations is *only* observed when the constituent is obligatory is challenged by Rossi *et al.* (2006), who observed an ELAN in L2 learners in response to word category violations in optional constituents. Rossi *et al.* tested a group of native Italian late learners of German and a group of native German late learners of Italian. L2 learners in both groups had been exposed after the age of 10, but differed in their proficiency: half of the speakers in each group were moderately proficient, and the other half were highly proficient. The learners listened to active voice sentences containing a word category violation in which the verb directly followed the preposition so that the noun expected in the prepositional phrase was absent (e.g. *\*Der Junge im singt ein Lied* [*\*‘The boy in the sings a song’*]). Both moderately proficient and highly proficient learners showed an ELAN between 100–250ms and an additional

negativity (reflecting reference-related information; see Rossi *et al.*, 2006), followed by a P600. This pattern is similar to that observed in native speakers. The moderately proficient learners showed the same three-component ERP pattern, but there was a quantitative difference in that their P600 was delayed and was of lower amplitude.

Rossi *et al.* (2006) thus found an ELAN in L2 learners, even with moderate proficiency in their L2, unlike the other studies reviewed above. One possible reason for the discrepancy is that Hahne (2001) and Hahne and Friederici (2001)'s sentences were in passive voice, whereas Rossi *et al.*'s sentences were in active voice, and processing simple active sentences may be easier than passives and/or the learners may have had greater exposure to active sentences. The presence of the ELAN in L2 learners with high and moderate proficiency levels suggests that with simple sentences, early automatic structure building processing can occur in L2 learners, even with only moderate proficiency. Rossi *et al.*'s findings also suggest that an ELAN can be elicited in response to word category violations in syntactically optional constituents, which seems inconsistent with Isel's proposal. Still, L2 learners may generate higher expectancies with respect to incoming word category in active sentences in which the noun is missing in the prepositional phrase (and violations in such constituents will thus be easier to detect) than in Isel's more complex sentences in which the violation occurred in the predicative prepositional phrase of a subject-modifying relative clause (e.g. \**Le chauffeur qui est dans la dort* [\*'The driver who is in the is sleeping']). It may thus be that ELAN in L2 learners occurs in phrase structure violations in constituents that enable L2 learners to generate high expectancies with regard to the expected word category, which is obviously the case in obligatory constituents and which may be the case in specific optional constituents.

To conclude, the presence of an ELAN in response to phrase structure violations in obligatory constituents (Isel, 2007) or in simple active voice sentences (Rossi *et al.* 2006), and the absence of an ELAN in response to phrase structure violations occurring in optional constituents (Weber-Fox and Neville, 1996; Hahne, 2001; Hahne and Friederici, 2001) suggest that early, automatic phrase-building processes in L2 learners may depend on the syntactic status of the sentence constituent and the parser's concomitant high expectancies with regard to incoming word category. Additionally, both Isel and Rossi *et al.* tested L2 learners whose native language and second language were both Indo-European

languages with close similarities in phrase structure (at least the violated structures were violations in both languages). In contrast, the studies that did not find an ELAN in L2 learners tested L2 learners whose native languages were markedly different from their L2, both in script and in sentence structure. In line with the Competition Model, if L2 learners' L1 and L2 are dissimilar, the L2 cues will be weaker than the L1 cues, leading to competition. During L2 sentence processing, such learners will thus not generate high expectations with respect to incoming word category information. Hence, ERP responses reflecting early, automatic structure building processes may not be observed in these learners during L2 sentence processing. In relation to proficiency, highly proficient learners sometimes demonstrate the same pattern of activity as native speakers (Rossi *et al.*, 2006). Isel also found a qualitatively similar pattern for natives and proficient learners, but with some quantitative differences. In some cases, the influence of proficiency is more difficult to discern because multiple factors differed between conditions. For example, in Weber-Fox and Neville (1996) proficiency decreased with increased age of exposure, and in the comparison between Hahne (2001) and Hahne and Friederici (2001) both L1/L2 similarity and proficiency level differed.

In all, these studies reveal several important and complex factors in ERP studies on grammatical processing in L2, including the similarity and dissimilarity of L1 and L2 syntax, the precise nature of the violated structure, and L2 proficiency. The studies discussed in this section provide an important breeding-ground for future research. In particular, there is a need for studies that take several of these factors into account to allow determination of the independent contributions of these factors to L2 processing.

*b Violations of morphosyntax:* In this section, we discuss studies that examined violations of morphology, and in particular violations of participial inflection, noun plurals, subject–verb agreement, determiner number and gender agreement, and auxiliary omission (see Table 1).

Hahne *et al.* (2006) examined processing of morphological inflection in native Russian speakers highly proficient in German. They tested verb participial inflection and noun plurals, which are both notably different in German and Russian. Two types of participial inflection violations were tested: regularizations (regular *-t* added to irregular verbs, \**gelauf*t vs. *gelaufen* ['walked']) and irregularizations (irregular *-n* added to

regular verbs, \**getanzen* vs. *getanzt* ['danced']). Similarly, two types of noun plural violations were tested: regularizations (regular -s added to nouns requiring -n, \**Vases* vs. *Vasen* ['vases']) and irregularizations (-n added to nouns requiring -s, \**Waggonen* vs. *Waggons* ['wagons']).

For participial forms, the L2 learners responded differently to regularizations and irregularizations. Specifically, regularizations elicited a bilateral anterior negativity between 250–600 ms and a P600. By contrast, native German speakers showed a LAN in regularizations (Penke *et al.*, 1997), but not a P600. Irregularizations elicited an N400 in learners and in native speakers. For plural formation, violations again revealed different responses for regularizations and irregularizations. The regularizations yielded no anterior negativity (unlike what has been found in native German speakers; Hahne *et al.*, 2006), although they did elicit a P600 (like in native speakers). Finally, as was found in the verb irregular forms, the noun plural irregularizations elicited an N400-like waveform, as had been found in native German speakers.

Hahne *et al.* (2006) concluded that L2 learners employ not only early automatic word-internal morphological processes in parsing verb-regularization violations (anterior negativity between 250–600 ms), but also later processes related to repair and effortful syntactic integration (P600), during which the participle is integrated with the rest of the sentence. In native speakers, participial regularization elicited anterior negativity, but no P600. This suggests that Hahne *et al.*'s L2 learners employed regular rules of verb inflection during early morphological processing, like native speakers. But, unlike with native speakers, regularization errors not only affect early morphological processing, but also later more controlled sentence-level integration and repair processes (P600).

Regularization errors of noun plurals, on the other hand, elicited no anterior negativity in the L2 learners, only a P600. This suggests that L2 learners employ later stage sentence-level repair and integration processes in both verb and noun plural regularization errors, but they employ earlier automatic word-internal morphological processes only in verb participles. To explain this result, Hahne *et al.* point to easier learning for participial inflection. German has many noun plural forms, and only a small proportion of nouns are inflected according to a default rule. In the verb participial system, by contrast, many verbs are inflected according to a regular rule, and fewer irregulars exist. L2 learners of German can therefore attain native-like proficiency earlier in verb participial inflection than in noun plural inflection. Accordingly, the

L2 learners' ERP patterns involved in parsing verb participles may be more automatized and more native-like than that of noun plurals.

The morphosyntactic rule tested most frequently in L2 learners is subject–verb agreement. This type of agreement was tested by Rossi *et al.* (2006) in addition to word category violations. They tested native Italian late L2 learners of German and native German late L2 learners of Italian; half of the speakers in each group were relatively less proficient in their L2 and the other half were highly proficient. The learners listened to sentences with third person singular subject–verb agreement violations and their correct counterparts. The higher proficiency learners of both German and Italian showed the biphasic pattern that had also been observed in native speakers of the two languages: a LAN (reflecting the detection of the morphosyntactic agreement error), followed by a P600 (reflecting a later re-analysis and integration process). In contrast, the moderately proficient German and Italian learners showed no LAN, only a P600. Moreover, compared to the highly proficient learners, the P600 in the moderately proficient learners had a delayed onset (also reported in Hahne, 2001) and a reduced amplitude.

Many studies on morphosyntactic processing have examined L2 learners whose L1 and L2 were Indo-European languages. By contrast, Chen *et al.* (2007) and Ojima *et al.* (2005) examined subject–verb agreement in Chinese and Japanese native speakers learning English. Chinese uses virtually no grammatical morphology to mark number, gender or case, so the learning of an L2 with these morphological markers may be fundamentally different for Chinese native language speakers. Likewise, Japanese has a system of agreement, but has no features such as person or number that induce subject–verb agreement. Chen *et al.* tested students at Chinese universities who began learning English at age 12. They were moderately proficient in English, as evidenced by their self-rated proficiency, an independent proficiency test and their grammaticality judgement performance on the critical sentences. The learners, as well as a group of native English speakers, read English sentences that contained third person singular subject–verb agreement violations and their correct counterparts. The native English speakers showed a biphasic LAN-P600. The L2 learners' ERPs were qualitatively different: they showed a late anterior-central negativity in the 500–700ms window (which the authors tentatively labelled an N600). So, unlike Rossi *et al.*'s German/Italian L2 learners of Italian/German, these Chinese learners of L2 English did not show a LAN or P600 in response to subject–verb agreement violations.

Ojima *et al.* (2005) also examined ERP responses to visually presented sentences with subject–verb agreement violations in L2 English. They tested native English speakers and native Japanese speakers who started to learn English in Japan between ages 10–12, and who had attained either intermediate or high proficiency as assessed by an independent formal test and self-ratings of proficiency. Subject–verb agreement violations elicited an enhanced left-lateralized negativity between 350–550 ms in native speakers and in highly proficient learners, but not in moderately proficient learners. This LAN was followed by a P600 in the native speakers, but not in either group of learners. Hence, the native speakers showed the biphasic LAN-P600 pattern, the highly proficient learners showed a LAN but no P600, and the moderately proficient learners showed none of the ERP components typically associated with syntactic processing.

In sum, together the subject–verb agreement studies indicate that morphosyntactic processing in native Italian, German or English speakers entails automatic word-internal morphological processes (LAN) followed by more controlled syntactic (re)analysis or integration (P600). When processing in L2, learners demonstrate this biphasic pattern only if L1 and L2 are similar and if they are highly proficient in L2 (see Rossi *et al.*, 2006). If L1 and L2 are dissimilar, only highly proficient learners show ERP components also observed in native speakers (Ojima *et al.*, 2005). However, if the L1 and L2 are dissimilar and the L2 learners are only moderately proficient, subject–verb agreement violations do not elicit syntax-related components (Chen *et al.*, 2007).

These ERP studies thus provide evidence for different ERP signatures of (morpho)syntactic processing in highly and moderately proficient L2 learners. Additional evidence suggesting that proficiency level is a more important factor than age of first exposure to L2 comes from fMRI and PET studies that have examined which brain regions are involved in L1/L2 processing. For low proficiency learners, different brain regions are recruited for L1 and L2 processing, whereas highly proficient learners recruit similar neural substrates in L1 and L2. Such a difference between low and high proficiency learners was even found when age of acquisition was controlled for (for more details, see Abutalebi and Green, 2007).

The studies discussed so far examined syntactic processing in late L2 learners who had achieved at least an intermediate level of L2 proficiency. Moreover, the majority of L2 learners began learning L2 around puberty. Recent studies by Tokowicz and MacWhinney (2005)

and Osterhout *et al.* (2006) differ from these studies in that they tested adult L2 learners in the earliest stages of L2 learning. Tokowicz and MacWhinney tested native English speakers who were in the first four semesters of university-level Spanish. Participants read L2 sentences with and without syntactic violations. The stimuli included three critical constructions that varied in the extent to which they would be similar when translated word for word between Spanish and English. One of the constructions was formed similarly in Spanish and English (auxiliary omission), one was formed differently in Spanish and English (determiner number agreement) and one was unique to Spanish (determiner gender agreement). The learners demonstrated P600s in response to violations for the similar and unique constructions but not for the different construction. Tokowicz and MacWhinney took these findings as evidence that the similarity between L1 and L2 can influence the likelihood that implicit processing will take place and described their findings in the Competition Model (MacWhinney, 1997) framework, using the reliability and validity of cues in L2 to explain why participants showed sensitivity for the unique condition.

A similar finding regarding L1/L2 similarity was reported by Osterhout *et al.* (2006) for native English speakers in the first year of university French instruction. The learners were tested after 1, 4 and 8 months of instruction on three types of violations: semantic violations, subject–verb (person) agreement which is phonologically realized in French, and determiner number agreement which is not phonologically realized and differs in French and English. Native French speakers demonstrated an N400 in response to the semantic anomalies and a P600 in response to both kinds of syntactic anomaly. A qualitatively different pattern was observed in the L2 learners, at least in the earliest phase of L2 learning. After one month of instruction, the French learners demonstrated an N400 in response to both the semantic and the subject–verb agreement violations, and were not sensitive to violations of determiner number agreement. Thus, at this point in learning, some syntactic violations were treated as word-level violations rather than as rule-based knowledge (compare Hahne *et al.*, 2006), who observed an N400 in response to irregularization errors in verb participial and noun plural inflections. After four months of instruction, the learners demonstrated a P600 in response to the subject–verb agreement violations. However, even after eight months of study they did not demonstrate reliable sensitivity to the determiner number agreement violations.

Although it is not possible to determine whether L1/L2 dissimilarity or a lack of phonological realization was responsible for the lack of sensitivity to determiner number violations, these findings are similar to those of Tokowicz and MacWhinney (2005) in the condition that differed between English and Spanish; they also used determiner number agreement and the learners did not demonstrate sensitivity even though this type of agreement is phonologically realized in Spanish. Thus, the dissimilarity between languages seems sufficient to produce this pattern of results. Taken together, the results of the two studies with beginning learners highlight the importance of L1/L2 similarity in the processing of L2 morphosyntax.

In the remainder of this article, we:

- review the implications of the ERP studies in relation to L1/L2 similarity, proficiency and concomitant expectancies;
- note some methodological issues in the study of L2 proficiency development;
- point out implications for theoretical perspectives on L2 sentence processing; and
- provide some ideas for a future research agenda.

### 3 *Can late L2 learners process an L2 in a native-like way?*

The available evidence indicates that the answer to whether late L2 learners process L2 in a native-like way is different for semantic processing and for (morpho)syntactic processing in the L2. L2 learners' ERP responses to semantically anomalous sentences resemble those of native speakers. Like native speakers, L2 learners at different levels of proficiency, even in the earliest phases of L2 learning, show an N400 in response to semantic anomalies (e.g. Weber-Fox and Neville, 1996; Hahne and Friederici, 2001; Ojima *et al.* 2005; Isel, 2007).

In contrast, L2 learners' ERP patterns associated with syntactic processing sometimes differed from those observed in native speakers. As shown in Section III.2a, phrase structure violations elicited a biphasic ERP pattern in native speakers: an ELAN (reflecting first-pass, early automatic syntactic structure building processes) followed by a P600 (reflecting later controlled processes of re-analysis and repair at the sentence level). In L2 learners, ELANs were not observed, or, as qualified in later studies, were observed only when phrase structure violations

occurred in obligatory constituents or in simple active voice sentences. However, it is difficult to disentangle these factors from proficiency because often multiple factors varied between studies. Nonetheless, this overall pattern suggests that early automatic structure-building processes in learners depend on the specific syntactic structure that is tested, and are more likely to occur when learners can generate high expectancies with regard to the incoming word category. The amount of overlap between syntactic structures in L1 and L2 is also important, in line with the Competition Model. If syntactic structures are dissimilar in L1 and L2, L2 cues will have lower strength for the learner. Hence, during L2 processing, learners will not generate high expectations with respect to incoming word category information, and processes related to early automatic structure-building may not occur.

The available evidence indicates that L2 learners' ERP responses to violations in L2 morphosyntax also sometimes differ from those of native speakers. The differences, again, appear to be related to L1/L2 similarity and to L2 proficiency. The native speaker data from most of these studies (but see Osterhout *et al.*, 2006) showed a biphasic ERP pattern in response to morphosyntactic violations: a LAN (indexing automatic word-internal morphological processes) followed by a P600 (indexing more controlled processes of syntactic (re)analysis or integration). In L2 sentence processing, learners show a native-like biphasic pattern only when the structure of the L2 is similar to their L1, and when they are proficient in their L2. When the L2 structures are dissimilar from those in L1, only highly proficient L2 learners show ERP components also observed in native speakers, in particular an anterior negativity that reflects automatic word-internal morphological processing. If L1 and L2 are dissimilar and the learners are only moderately proficient in L2, violations of morphosyntax do not appear to elicit syntax-related ERP components. Finally, studies that examined adult L2 learners in the earliest stages of L2 learning also indicate that some (morpho) syntactic violations are treated as word-level violations eliciting N400s rather than syntax-related components (e.g. Osterhout *et al.*, 2006), and that beginning L2 learners show syntax-related ERP components only in response to violations in L2 structures that are similar to their L1 (or unique; see Tokowicz and MacWhinney, 2005).

Crucially, there are some studies in which earlier components (ELAN, LAN) are not observed even in native speakers (e.g. Ainsworth-Darnell *et al.*, 1998; Osterhout *et al.* 2006). Thus, the likelihood of observing earlier components seems to depend on the type of violation and

perhaps also specific experimental parameters (mode of presentation, timing, etc.). See Osterhout *et al.* (2006: 207; 226) for additional issues with using these earlier components to study L2 learning.

To conclude, whether L2 learners show native-like ERP signatures of (morpho)syntactic processing depends on several, partly interrelated factors: similarity of (morpho)syntactic structures across languages, the ability to generate high expectancies with regard to particular violations, and L2 proficiency level. The review of studies in this article reveals that the most pronounced qualitative differences are found between moderately proficient learners and highly proficient L2 learners/native speakers. L2 sentence processing in moderately proficient learners is often characterized by the absence of components indexing early automatic structure building and word-internal morphological processes, and later processes of (re)analysis and repair (see Table 1). In most studies, the same sentence materials did elicit these components in proficient L2 learners and native speakers.

Both qualitative and quantitative differences in the ERP components associated with syntactic processing appear to be related to L2 proficiency. Rossi *et al.* (2006) observed a delayed P600 onset and reduced amplitude in response to phrase structure violations and subject–verb agreement violations in moderately proficient learners, but not in highly proficient learners. However, Hahne (2001) also observed a delayed P600 peak onset in response to phrase structure violations in proficient L2 learners. Obviously more research is needed to elucidate the, sometimes subtle, quantitative differences between learners at different proficiency levels.

#### **IV Methodological note: approaches in the study of L2 proficiency development**

The review of ERP studies shows that L2 proficiency is an important factor in the ERP patterns that are obtained, and in the differences in the underlying cognitive processes these ERP components are assumed to reflect. A recurrent issue in studies on L2 learning is that it is very difficult to compare the influence of L2 proficiency on L2 processing across studies. In addition to variations in L2 proficiency among the L2 learners, studies typically also vary on a wealth of other variables (languages involved and the (dis)similarity of structures across languages, specific syntactic structures studied, etc.), which potentially limits comparability across studies. Moreover, in quite a few studies proficiency is confounded with age of first exposure.

An elegant approach to gain more insight into the role of L2 proficiency is to examine two groups of bilinguals from the same population with the same age of first exposure to the L2, but who achieved different levels of proficiency in the course of their L2 learning. This approach was taken by Ojima *et al.* (2005), leading to clearly interpretable results on how L2 proficiency modulates ERPs associated with syntactic processing in L2. A second valuable approach to the role of proficiency in the course of individual development is via longitudinal studies of L2 learning (for an extensive discussion, see Osterhout *et al.*, 2006).

A third approach in which L2 proficiency – as well as L1/L2 similarity – can be systematically manipulated is via ‘miniature studies’ (Friederici *et al.*, 2002; Mueller *et al.*, 2005). These researchers used miniature natural or artificial trained languages to explore whether adult L2 learners can become native-like. One advantage of this approach is that participants can become highly proficient. Furthermore, proficiency can be manipulated explicitly as can L1/L2 similarity. Friederici *et al.* (2002) taught a group of native German speakers a miniature artificial language called BROCANTO. The experimental group was extensively trained on BROCANTO and became highly proficient. The control group was given only vocabulary training to isolate the effects of syntactic knowledge. Sentences with or without syntactic word category violations were auditorily presented to the learners. The experimental group showed a bilaterally distributed early anterior negativity and a P600, whereas the control group showed neither. The highly proficient learners thus show the biphasic ERP pattern of syntactic processing, assumed to reflect early, automatic first-pass structure building processes followed by more controlled second-pass processes. Furthermore, a subsequent analysis of only the rules that could not have been transferred from L1 revealed that the finding of biphasic syntactic sensitivity was due to learning and not L1 transfer. The findings thus show that a late-learned language – provided that the learners achieve high proficiency – can be processed in a native-like way (but, for a different pattern of results, see Mueller *et al.*, 2005).

## **V Implications for theoretical perspectives on L2 sentence processing**

Earlier in this article we outlined several theoretical perspectives that have been used as theoretical background in ERP studies on L2 (morpho)

syntactic processing: the critical period hypothesis, the age-related decline framework, the Competition Model and the declarative/procedural model. To what extent does the available ERP evidence support these theoretical perspectives?

The critical period hypothesis maintains that there is a biologically determined critical period in language acquisition, which implies that late learners of an L2 cannot achieve native-like proficiency. Several lines of evidence seriously challenge the tenability of this position. Weber-Fox and Neville (1996) found that all L2 learners' ERP responses to syntactic violations, irrespective of the age of first exposure to L2 (from 1–3 to >16 years), were different from those of native speakers. This pattern is also not consistent with a strict age-related decline. Moreover, quite a number of studies observed only quantitative but no qualitative differences in ERP components of L2 learners and native speakers. Perhaps more challenging evidence is provided by Ojima *et al.* (2005) who tested L2 learners with different proficiency levels, but the same age of first exposure, and found that the highly and moderately proficient learners' ERP responses to morphosyntactic violations were qualitatively and quantitatively different. This indicates that L2 proficiency level is more decisive than mere age of first exposure. This conclusion is further supported by brain imaging studies, which also indicate that L2 proficiency level is a more important factor than age of exposure (see Abutalebi and Green, 2007). Probably the most compelling line of counter-evidence comes from studies in which adults have been exposed to an L2 for a relatively short period of time. Osterhout *et al.* (2006) tested adult L2 learners in the beginning stages of L2 learning and showed that their ERP patterns in response to at least some structures were not different from those of native speakers.

The Competition Model emphasizes the possible transfer of syntactic structures from L1 to L2. The reliability and availability in the L2 input are crucial, and as L1 and L2 structures are more similar, the L2 will be easier to learn because L2 learners benefit from cross-language transfer. The ERP studies with beginning L2 learners – as well as those with moderately proficient and highly proficient learners – indeed indicate that L1/L2 similarity affects syntactic and morphosyntactic processing in L2.

The declarative/procedural model distinguishes declarative and procedural processes that can be indexed by early and later ERP components. It is proposed that L1 syntax relies on procedural processing but that this kind of processing is unavailable to L2 learners until very

high proficiency has been reached. Similarly, Hahne and Friederici and colleagues have proposed that earlier, automatic structure-building or word-internal morphological processes are less automatized in L2 learners, and are thus more likely to be different from native speakers. Later, more controlled processes like syntactic re-analysis and repair are more likely to resemble those of native speakers. Many studies indeed found that differences in ERP patterns between L2 learners and native speakers are most pronounced in the earlier syntax-related components, and report an absence of components that are assumed to index early automatic structure-building processes (ELAN) or word-internal morphological processes (LAN) in L2 learners (see Table 1). However, other studies demonstrated only quantitative differences between less and more proficient learners, and some even found native-like patterns early in learning. These findings are not entirely consistent with the view that learners only become native-like after very high proficiency has been reached.

The available evidence suggests that the absence or presence of ELAN or LAN in L2 learners varies with several factors:

- the type of syntactic structure and the expectancies L2 learners can generate with respect to violations of this structure;
- the degree of L1/L2 syntactic structure similarity; and
- L2 proficiency level.

These factors are likely to be interrelated. For example, L2 learners may not be able to generate high expectations with regard to certain syntactic structures, because the L2 syntactic structures are dissimilar from L1. This pattern, in turn, will change when L2 learners become more proficient, and progress from using L1 cues to using more native-like cues to comprehend the L2.

## **VI Concluding remarks**

Can the monitoring of brain activity using the ERP technique contribute to our knowledge of second language learning, one of the questions *Science* formulated for the next 25 years? One of the benefits of the ERP technique – as the studies discussed in this article indicate – is that ERPs can be used to elucidate subtle differences in syntactic processing that are difficult to detect, or that might have remained undetected, with behavioural studies. For example, ERP measures provide

a possible way to discern earlier automatic and later more controlled processes – and the timing of these processes – as L2 sentence comprehension unfolds over time. ERP measures can also detect subtle quantitative and qualitative differences in the timing and degree of neural activation during syntactic processing between L2 learners and native speakers, or between L2 learners at different proficiency levels; these are differences that would remain undetected in behavioural measures like the frequently used grammaticality judgement task. ERP studies have also allowed us to conclude that learning has taken place much earlier than would have been thought based on behavioural performance alone. This is because in several studies behavioural performance was at chance when ERP responses demonstrated significant sensitivity to violations of syntax (e.g. Tokowicz and MacWhinney, 2005; see also McLaughlin *et al.*, 2004). In particular, Tokowicz and MacWhinney (2005) found that L2 learners demonstrated the strongest ERP effects (P600) for the construction for which they had the lowest accuracy (unique to L2: grammatical-gender agreement). This and similar findings indicate that behavioural responses sometimes diverge from brain responses, suggesting the importance of ERPs as a converging method in the study of L2 learning.

The relatively few studies on ERPs and grammatical processing in late L2 learners, which all appeared in a 10-year time window, have provided important insights into the electrophysiological correlates for grammatical processing in L2 learners. These studies indicate that patterns of neural activity in the brain during syntactic and morpho-syntactic processing can be modulated by various, possibly interrelated factors, including:

- the similarity or dissimilarity of syntactic structures in L1 and L2;
- the exact nature of the syntactic structure L2 learners seek to comprehend, and the concomitant expectancies they can generate with regard to violations in this structure; and
- the L2 learners' level of proficiency.

These are important insights, and the factors form a valuable breeding ground for future research to address the many issues that still need to be clarified. Some of these issues are methodological (e.g. whether or not only correct trials are included in the ERP analyses), whereas other issues are more theoretical or groundbreaking. An example of the latter is L2 learning in children. Late L2 learners – who in fact constitute a

large portion of the adult bilingual population – often learn their L2 around puberty, and often in a classroom. Indeed, the majority of the late L2 learners who participated in the studies reviewed here had learned their L2 in such a setting. Remarkably, the few studies that examined ERPs in the earliest stages of L2 learning all examined *adult* beginning L2 learners. We know little, if anything, about the neural correlates of L2 learning in *child* L2 learners, which is a topic we are currently addressing in our lab (e.g. Brenders *et al.*, 2007). Recording ERPs in child L2 learners is a challenging enterprise, because, as discussed above, ERP components of both children and L2 learners tend to diverge from the known characteristics of ERP components based on research with adult native speakers. But if we want to provide a full answer to the question posed in the 125th anniversary edition of *Science* (Kennedy and Norman, 2005), the study of child late L2 learners should also be included in our research agenda for the coming 25 years.

## VII References

- Abutalebi, J. and Green, D.** 2007: Bilingual language production: the neurocognition of language representation and control. *Journal of Neuro-linguistics* 20, 242–75.
- Ainsworth-Darnell, K., Shulman, R. and Boland, J.** 1998: Dissociating brain responses to syntactic and semantic anomalies: evidence from event-related brain potentials. *Journal of Memory and Language* 38, 112–30.
- Birdsong, D.** 1999: *Second language acquisition and the critical period hypothesis*. Mahwah, NJ: Lawrence Erlbaum.
- Brenders, P., van Hell, J.G. and Dijkstra, T.** 2007: The development of lexical and conceptual processing in classroom L2 learners: ERP and behavioural evidence. *Paper presented at the 6th International Symposium on Bilingualism*. Hamburg.
- Chen, L., Shu, H., Liu, Y., Zhao, J. and Li, P.** 2007: ERP signatures of subject–verb agreement in L2 learning. *Bilingualism: Language and Cognition* 10, 161–74.
- Fabiani, M., Gratton, G. and Coles, M.G.H.** 2000: Event-related brain potentials: methods, theory, and applications. In Cacioppo, J.T., Tassinary, L.G. and Berntson, G.G., editors, *Handbook of psychophysiology*. New York: Cambridge University Press, 53–84.
- Friederici, A.D.** 2002: Towards a neural basis of auditory sentence processing. *Trends in Cognitive Science* 6, 78–84.
- 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 Science (PNAS)* 99, 529–34.

- Hagoort, P., Brown, C.M. and Groothusen, J.** 1993: The syntactic positive shift (SPS) as an ERP measure of syntactic processing. *Language and Cognitive Processes* 8, 439–83.
- Hagoort, P., Hald, L., Bastiaansen, M. and Petersson, K.M.,** 2004: Integration of word meaning and world knowledge in language comprehension. *Science* 304, 438–41.
- Hahne, A.** 2001: What's different in second-language processing? Evidence from event-related brain potentials. *Journal of Psycholinguistic 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, 193–204.
- 2001: Processing a second language: late learners: comprehension mechanisms as revealed by event-related brain potentials. *Bilingualism: Language and Cognition* 4, 123–41.
- Hahne, A., Mueller, J. and Clahsen, H.** 2006: Morphological processing in a second language: behavioural and event-related potential evidence for storage and decomposition. *Journal of Cognitive Neuroscience* 18, 121–34.
- Handy, T.C.** 2005: *Event-related potentials: a methods handbook*. Cambridge, MA: MIT Press.
- 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.
- Isel, F.** 2007: Syntactic and referential processes in second-language learners: event-related brain potential evidence. *Neuroreport* 18, 1885–89.
- Isel, F., Hahne, A., Maess, B. and Friederici, A.D.** 2007: Neurodynamics of sentence interpretation: ERP evidence from French. *Biological Psychology* 74, 337–46.
- Kaan, E., Harris, A., Gibson, E. and Holcomb, P.J.** 2000: The P600 as an index of integration difficulty. *Language and Cognitive Processes* 15, 159–201.
- Kennedy, D. and Norman, C.** 2005: 125 questions: what don't we know? *Science* 125 (volume 309), 78–102.
- Kutas, M. and Hillyard, S.A.** 1980: Reading senseless sentences: brain potentials reflect semantic incongruity. *Science* 207 (4427), 203–05.
- Lenneberg, E.** 1967: *Biological foundations of language*. New York: Wiley.
- Luck, S.J.** 2005: *An introduction to the event-related potential technique*. Cambridge, MA: MIT Press.
- MacWhinney, B.** 1997: Second language acquisition and the Competition Model. In de Groot, A.M.B. and Kroll, J.F., editors, *Tutorials in bilingualism: psycholinguistic perspectives*. Mahwah, NJ: Lawrence Erlbaum, 113–42.
- McDonald, J.L.** 1987: Sentence interpretation in bilingual speakers of English and Dutch. *Applied Psycholinguistics* 8, 379–414.

- McLaughlin, J., Osterhout, L. and Kim, A.** 2004: Neural correlates of second-language word learning: minimal instruction produces rapid change. *Nature Neuroscience* 7, 703–04.
- Moreno, E.M., Rodriguez-Fornells, A. and Laine, M.** 2008: Event-related potentials (ERPs) in the study of bilingual processing. *Journal of Neuro-linguistics* 21, 477–508.
- 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.
- Ojima, S., Nakata, H. and Kakigi, R.** 2005: An ERPs study of second language learning after childhood: effects of proficiency. *Journal of Cognitive Neuroscience* 17, 1212–28.
- Osterhout, L. and Holcomb, P.** 1992: Event-related brain potentials elicited by syntactic anomaly. *Journal of Memory and Language* 31, 785–806.
- Osterhout, L., McLaughlin, J., Pitkänen, 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, 199–203.
- Penke, M., Weyeerts, H., Gross, M., Zander, E., Münte, T.F. and Clahsen, H.** 1997: How the brain processes complex words: an event-related potential study of German verb inflection. *Cognitive Brain Research* 6, 37–52.
- 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.
- Tokowicz, N. and MacWhinney, B.** 2005: Implicit and explicit measures of sensitivity to violations in second language grammar: an event-related potential investigation. *Studies in Second Language Acquisition* 27, 173–204.
- Tokowicz, N. and Warren, T.** 2008: Quantification and statistics. In Wei, L. and Moyer, M., editors, *The Blackwell guide to research methods in bilingualism and multilingualism*. Malden, MA: Blackwell, 214–31.
- 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.
- van Hell, J.G. and Wittenman, M.J.** 2009: The neurocognition of switching between languages: a review of electrophysiological studies. In Isurin, L., Winford, D. and de Bot, K., editors, *Multidisciplinary approaches to code switching*. Amsterdam/Philadelphia: John Benjamins, 53–84.
- Weber-Fox, C.M. and Neville, H.J.** 1996: Maturation constraints on functional specializations for language processing: ERPs and behavioural evidence in bilingual speakers. *Journal of Cognitive Neuroscience* 8, 231–56.