

## Novice Learners, Longitudinal Designs, and Event-Related Potentials: A Means for Exploring the Neurocognition of Second Language Processing

*Lee Osterhout, Judith McLaughlin, and Ilona Pitkänen*  
*University of Washington*

*Cheryl French-Mestre Centre*  
*National de la Recherche Scientifique Université de Provence*

**Q1**

*Nicola Molinaro*  
*Università degli Studi di Padova*

Research on the neurobiology of second language (L2) learning has historically focused on localization questions and relied on cross-sectional designs. Here, we describe an alternative paradigm involving longitudinal studies of adult, novice learners who are progressing through an introductory sequence of classroom-based L2 instruction. The goal of this paradigm is to determine how much L2 exposure is needed before learners incorporate L2 knowledge into their online comprehension processes, as reflected in scalp-recorded event-related brain potentials. Our preliminary studies show that some, but not all, aspects of the

---

Lee Osterhout, Judith McLaughlin, and Ilona Pitkänen, Department of Psychology; Cheryl French-Mestre Centre, National de la Recherche Scientifique Université de Provence; Nicola Molinaro, Università degli Studi di Padova.

**Q2**

The authors thank Julia Herschensohn for her comments and collaboration. We received financial support from grants R01DC01947 and P30DC04661 from the National Institute on Deafness and Other Communication Disorders, National Institutes of Health.

Correspondence concerning this article should be addressed to Lee Osterhout, Department of Psychology, Box 351525, University of Washington, Seattle, WA 98195. Internet: [losterho@u.washington.edu](mailto:losterho@u.washington.edu)

200 *Novice Learners, Longitudinal Designs, ERPs*

L2 (including lexical and morphosyntactic aspects) are incorporated into the comprehension system after remarkably little L2 instruction. We discuss the benefits of this paradigm while acknowledging the limitations and potential difficulties associated with it.

One could argue that two historical facts have impeded progress in understanding how a second language (L2) is instantiated in a language learner's brain. The first historical fact is the emphasis on the "where" question (i.e., the question of which parts of the brain are involved in using the L2). Although the "where" question is a valid and interesting one, the focus on this question has led to neglect of other questions that might prove to be equally important. The second historical fact concerns how researchers have dealt with two sources of variability. The first type, variability among L2 learners, introduces potential confounding variables that are well understood but rarely adequately controlled for. The second type, variability within a single learner over time, is a potentially crucial source of evidence that has been largely ignored. Our goal in this article is to argue for an alternative paradigm for studying the neurocognition of L2 learning and usage, in which the focus is on how the processes underlying comprehension of the L2 change with increasing L2 exposure or proficiency. The paradigm minimizes the troublesome sources of between-learner variability while maximizing sensitivity to theoretically crucial within-learner variability over time.

Two primary methods have been used to identify the neural structures involved in using an L2. One method involves studying aphasic bilinguals. If a patient's facility with two languages can be damaged independently, then one reasonable interpretation is that the two languages are represented independently in the speaker's brain. A second method involves the use of neuroimaging techniques (e.g., functional magnetic resonance imaging [fMRI]) to contrast L1 and L2 localization within the brain. These tools provide a measure of the brain's metabolic activity in neurologically intact individuals and allow researchers to determine which areas of the brain are metabolically most active, given

some task. Evidence that different areas of the brain become activated when a bilingual uses her first (L1) or second language could, again, be interpreted to mean that different neural areas are involved in using the two languages.

However, nearly every pattern of language loss and recovery (loss of an L1 with preservation of an L2, equal loss of both languages, etc.) has been reported in bilingual aphasics. Similarly, whereas a few neuroimaging studies seem to indicate a (partial) neuroanatomical separation of the two languages (Kim, Relkin, Lee, & Hirsch, 1997; Perani et al., 1998), others seem to indicate a complete overlap in the brain areas activated by the two languages (Chee, Tan, & Thiel, 1999; Klein, Milner, Zatorre, Zhao, & Nikelski, 1999). It is reasonable to assume that the lack of consistency across experiments reflects, in part, uncontrolled subject variability (variability in learner motivation, the type and nature of L2 exposure, the age of L2 acquisition, etc.), a problem that has vexed much of the L2 literature (Grosjean, 1998). Furthermore, the range of inferences that these methods permit is quite limited; for example, language dysfunction might result from damage to white matter tracts that course through the area, rather than from damage to the gray matter at the lesion site. A complete overlap in fMRI activation patterns for an L1 and an L2 does not necessarily imply that the two languages are processed in exactly the same way; and a partial overlap in activation patterns gives us little idea of how processing of the two languages might differ.

Rather than focus on the “where” question, the paradigm we have in mind focuses on the processes that recognize words and derive the structure and meaning of sentences. More specifically, the paradigm is designed to help us learn more about how these processes (and the neural systems that underlie them) change over time, with increasing L2 exposure or proficiency. This paradigm minimizes between-subjects variability by longitudinally studying novice L2 learners who are progressing through their first years of classroom L2 instruction. In this design, the learners are highly similar in a priori L2 proficiency (having none)

202 *Novice Learners, Longitudinal Designs, ERPs*

and experience a highly similar learning environment. Because each learner acts as his or her own control, many of the subject variables that can potentially confound L2 studies have been eliminated within this design. By also including native speakers of the L2 in the study, we are able to compare the beginning and intermediate states of learning to the end state of native proficiency.

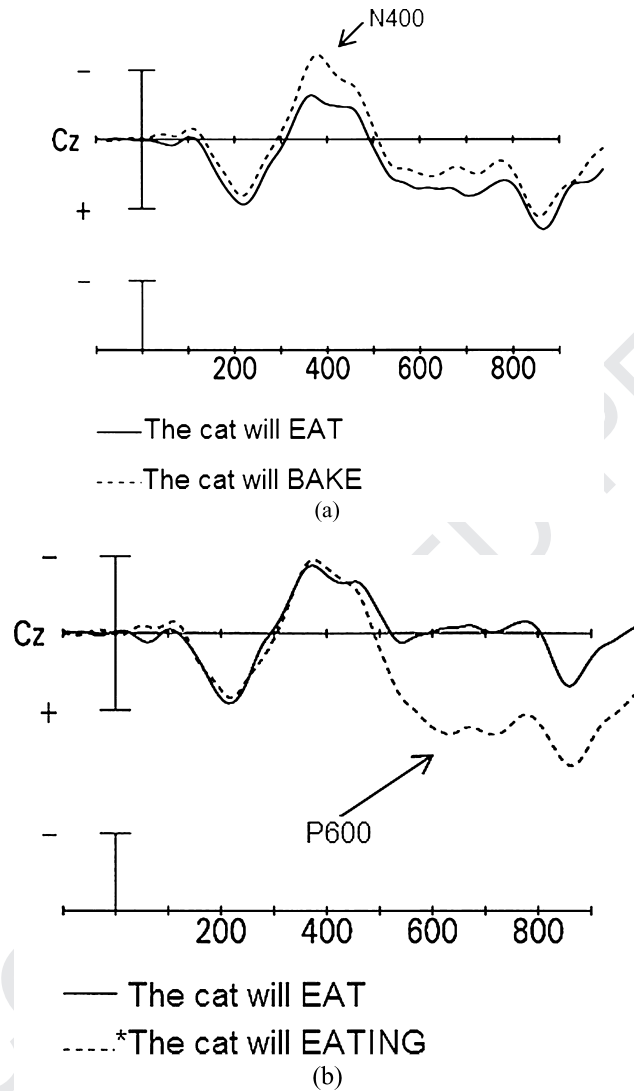
The longitudinal design also maximizes sensitivity to the within-subject variation that is the primary focus of our research program. Our goal is to identify changes in brain activity (and, in particular, activity that reflects online L2 processing) that accompany the earliest stages of L2 learning. We do so by recording event-related brain potentials (ERPs) from the scalp while learners read L2 words or sentences. ERPs provide an online, millisecond-by-millisecond record of the brain's electrical activity during language comprehension and, therefore (unlike fMRI and positron-emission tomography [PET]), have the necessary temporal resolution to isolate specific language comprehension processes in time. Furthermore, ERPs are multidimensional, varying in polarity, timing, morphology, and scalp distribution; this multidimensionality might in theory provide ERPs with a differential sensitivity to at least some of the processing steps underlying comprehension.

In fact, ERPs do seem to be differentially sensitive to events occurring at distinct levels of linguistic analysis (for a review, see Osterhout, McLaughlin, Kim, Greenwald, & Inoue, 2004). One ERP response, the N400 component, is sensitive to properties of words, both when they appear in isolation and when they appear in a linguistic context. A second component, the P600, is sensitive to the grammatical well-formedness of sentences. Although the precise cognitive and neural events underlying these effects are not known, their existence and particular sensitivities can be exploited to learn a great deal about language comprehension and (we hope to demonstrate here) L2 learning. We will demonstrate that although ERPs might not be tremendously informative with respect to the "where" question, they are, in some

instances, extremely useful for ascertaining *what* is happening and *when* it happens. The method's sensitivity to the "what" question follows from the fact that the language-relevant ERP effects are quite specific in their sensitivities, thereby permitting fairly specific inferences about what is happening at that moment in processing. The method's sensitivity to the "when" question encompasses two timescales: the small intervals of time (tens and hundreds of milliseconds) over which a word or sentence is processed, and the large intervals of time (weeks, months, and years) over which a person progresses from no competence with a language to increasing competence. In the next section, we briefly review what has been learned about the relationship between language processing and ERPs, discussing, in particular, problems that might come up when applied to L2 research. In the subsequent sections, we provide examples of how this knowledge, when combined with the above-described research paradigm, can be used to gain some new insights about what happens during L2 learning and processing and when it happens.

#### ERPs as Tools for Studying First and Second Language Processing

Kutas and colleagues (Kutas & Hillyard, 1980) were the first to demonstrate that semantically anomalous words (e.g., "He spread his warm bread with *socks*") elicit an increase in the amplitude of the N400 component, a negative-going wave that peaks at about 400 ms (Figure 1a). This result is observed regardless of the position of the anomalous word within the sentence or the modality of input (visual vs. auditory; Osterhout & Holcomb, 1993). Subsequent research has shown that the N400 amplitude to words in sentences is an inverse function of the semantic congruency between the target word and the preceding context, even when the word is not semantically anomalous (Kutas & Hillyard, 1984). The N400 amplitude is also sensitive to the strength of semantic priming in a word-pair lexical decision task (Bentin, McCarthy, & Wood, 1985). These results have been interpreted as indicating

204 *Novice Learners, Longitudinal Designs, ERPs*

*Figure 1.* (a) ERPs (recorded at central midline location Cz) elicited by semantically anomalous words (dashed line) and nonanomalous control words (solid line) in sentences such as *The cat will eat/bake the food*. (b) ERPs elicited by syntactically anomalous (dashed line) words and well-formed controls (solid line) in sentences such as *The cat will eat/eating the food*. Onset of the critical word is indicated by the vertical bar. Each hashmark represents 100 ms. The vertical calibration bar is 5  $\mu$ V. (Adapted from Osterhout & Nicol, 1999.)

that the N400 amplitude elicited by words is highly sensitive to the semantic relationship between the word and the preceding context. However, the N400 amplitude is not solely determined by the semantic fit between a word and its context. A variety of lexical properties affect the N400 amplitude, including lexicality (i.e., whether a letter string is a word in the language), word frequency, phonological priming, morphological content within a string (e.g., McKinnon, Allen, & Osterhout, 2003; McLaughlin, Osterhout, & Kim, 2004), and sequential probabilities concerning the likelihood of words occurring in succession (Kutas & Hillyard, 1984). A more accurate generalization is that the N400 amplitude reflects a combination of lexical and semantic/conceptual factors.

The sensitivity of the N400 component to lexico-semantic factors is strikingly contrasted with its *insensitivity* to syntactic factors. A great deal of research conducted over the past 25 years has shown that violations of syntactic rules do not robustly affect the N400 amplitude (Allen, Badecker, & Osterhout, 2003; Osterhout & Nicol, 1999). Instead, a disparate set of syntactic anomalies elicits a large positive wave in the ERP with an onset at about 500 ms and a duration of several hundred milliseconds [labeled the P600 effect by Osterhout & Holcomb (1992; Figure 1b)]. The P600 effect is highly sensitive to syntactic variables but insensitive to lexico-semantic variables (Allen et al.). Importantly, the N400/P600 dichotomy is not a function of particular combinations of stimuli, task, and language, but, instead, generalizes well across types of anomaly (including anomalies involving phrase structure, morphosyntax, and constituent movement), numerous stimulus and task conditions (e.g., rate of word presentation, subject's task), modalities (visual and auditory), and languages (including configurational languages like English and nonconfigurational languages such as Japanese). In some reports, syntactic anomalies have also elicited an anterior negativity within a window ranging from 150 to 500 ms, sometimes largest in amplitude over the left hemisphere (hence the label *left anterior negativity*, or LAN; Friederici, 1995; Neville, Nicol, Barss, Forster, & Garrett, 1991; Osterhout & Holcomb, 1992, 1993).

206 *Novice Learners, Longitudinal Designs, ERPs*

As is the case with any method, ERPs have their limitations. Although it is relatively easy to determine the antecedent conditions (i.e., the stimulus manipulations) that produce or modulate some effect, it is usually very difficult to identify the specific processes that are manifested by the effect; for example, even though the P600 effect is reliably elicited by violations of a syntactic rule, it does not necessarily follow that the P600 effect reflects specifically syntactic (or even linguistic) processes. This ambiguity exists because ERPs (just like neuroimaging techniques such as fMRI) are correlational in nature. Consequently, the P600 effect might reflect syntactic processes directly or it might reflect some process that is highly correlated with these syntactic processes or with the processes that respond to a syntactic anomaly. Deciding which of these possibilities is correct might prove to be an intractable problem.<sup>1</sup> Fortunately, the correlations between manipulations of lexical and semantic variables and changes in N400 amplitude and between the presence of a syntactic anomaly and the P600 effect are very robust. One can, therefore, reasonably infer that a linguistic anomaly was perceived to be semantically anomalous if it elicited an N400 effect or syntactically anomalous if it elicited a P600 effect.

A second limitation is that ERPs usually need to be averaged twice: once for a given subject and then over all subjects. This is required to achieve the necessary signal-to-noise ratio. When the subjects are native speakers of a language, there seems to be enough consistency (within and between subjects) in the ERP responses to linguistic stimuli to produce robust and reproducible results. However, a sample of L2 learners will almost always be more variable than a sample of L1 speakers. It seems quite likely that this variability will be expressed in the ERPs (as, e.g., greater variation in the latency or scalp distribution of particular ERP effects). If so, then the differences in variability limit the types of inference that are warranted by ERP data; for example, if an effect is present in the L1 group but absent in the L2 group, this does not necessarily mean that the process manifested by the effect is present in the L1 group but not in the



L2 group. This is because the effect might be present in both groups but obscured in the L2 group due to larger variability in the effect's timing and distribution. The increased variability would reduce the size of the effect, or eliminate it altogether, in the more variable population—even if it is present on individual trials, for individual subjects. An example of this problem is provided by a study reported by Hahne (2001). Hahne recorded ERPs to syntactic anomalies from native German speakers and from Russian L1 speakers who had learned German after the age of 10. Both groups of speakers showed a P600 effect to the anomalies, but only the native speakers also showed a LAN effect. Hahne interpreted this result as indicating that the German learners were missing the linguistic process manifested by the LAN effect. However, another possibility is that the greater variability in the L2 group simply obscured the anterior negativity, even if it was present on individual trials, in individual learners. This possibility is likeliest for effects that are small in amplitude and have a limited temporal extent (such as the LAN effect).<sup>2</sup>

The increased variability associated with L2 learners complicates the use of ERPs in another way. If the variability across subjects is large enough, the grand average might not reflect the response shown by any particular individual learner and might not generalize to any identifiable population; that is, in such cases, the “central tendency” represented by significant effects in the grand average might reflect only the accidental overlap of effects that were present in many learners in the sample, but it might not reflect larger and potentially more important effects that varied from learner to learner. If so, then comparisons among, for example, L2 learners who were exposed to the L2 at different ages might produce reliable differences among the groups but might not represent true effects of age on acquisition. Fortunately, the validity or invalidity of such results would become clear when the results prove difficult to replicate. Unfortunately, this problem could easily lead to the publication of many unreplicable and irreconcilable results.

## 208 *Novice Learners, Longitudinal Designs, ERPs*

Event-related brain potentials therefore offer a great deal of promise for learning more about the acquisition and use of an L2, but also carry with them certain risks. Our goal has been to reap the benefits of ERPs while minimizing our exposure to the risks. To that end, we have combined ERPs with longitudinal studies of novice language learners. Our questions have included the following: How quickly is L2 knowledge about words and sentences incorporated into the online, “real-time” processing system? What variables influence this rate? How quickly do learners’ brain responses to L2 words and sentences begin to approximate their responses to L1 words and sentences? What is the relationship between brain-based manifestations of learning and a person’s overt behavior? This approach might also usefully investigate some of the standard questions within the field of L2 research. The following are some examples: What influence does the learner’s L1 have on acquisition of the L2? What role does a correspondence (or lack of correspondence) between phonology and morphology play in the acquisition of a second language? Does L2 learning progress in a continuous manner, or are there discrete stages of learning? Although these questions are not directly about the cognitive neuroscience of L2, we hope to show that the tools of cognitive neuroscience are particularly well suited for providing answers to them. Also, because ERPs directly reflect brain activity, the results of these studies sometimes shed light on the cognitive neuroscience of L2s even when that is not the explicit goal.

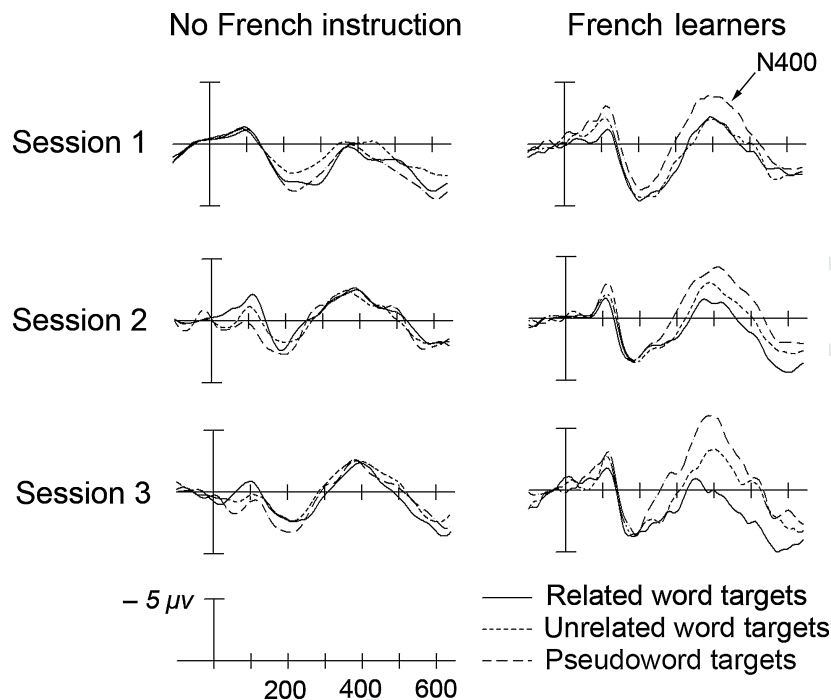
### ERPs and L2 Word Learning

Research on bilingual lexical processing has primarily focused on the organization of the bilingual lexicon (Kroll & Sunderman, 2003). Most theorists assume that the representation of words is divided into two levels: a word-form (lexical) level and a word-meaning (conceptual) level. The question of interest has been the degree of independence of the two lexicons at

these two levels (Kroll & Sunderman; Potter, So, Von Eckhardt, & Feldman, 1984).

However, little is known about how L2 lexical and conceptual representations develop during L2 acquisition (McLaughlin, 1998), and almost nothing is known about the neurobiological correlates of this developmental process (for one recent exception, see Raboyeau, Marie, Balduyck, Gros, Demonet, & Cardebat, 2004). In a study conducted in our laboratory, we examined word learning by measuring learning-related changes to the N400 component of the ERP (McLaughlin, Osterhout, & Kim, 2004). As noted earlier, N400 is sensitive to both lexical status (whether or not a particular wordlike form is part of the language) and word meaning (Bentin, 1987; Kutas & Hillyard, 1980). For native speakers of a given language, the N400 amplitude is largest for pronounceable, orthographically legal nonwords (pseudowords; e.g., *flirth*), intermediate for words preceded by a semantically unrelated context, and smallest for words preceded by a semantically related context. Our goal was to determine how much L2 exposure is needed before the learner's brain responses to L2 words and nonwords resemble that of a native speaker. Such results might reveal the rate at which French learners acquire information about word forms and word meanings.

Our participants were English-speaking university students progressing through their first year of classroom French instruction and members of a control group who had never received any French instruction. None of the students had significant exposure to French prior to attending class. The stimuli were prime-target pairs of letter strings. Some of the pairs contained two French words that were either semantically related (e.g., *chien-chat* "dog-cat") or unrelated (*maison-soif* "house-thirst"). Other pairs contained a target that was a pronounceable French pseudoword (*mot-naisier*). ERPs to the targets were of interest. The French learners were tested three times in a longitudinal design: once near the beginning of their French instruction (after ~14 hr on instruction), once near the middle (~60 hr), and once

210 *Novice Learners, Longitudinal Designs, ERPs*

*Figure 2.* ERPs from the No Instruction (left panel) and French Instruction (right panel) subjects, recorded at three successive longitudinal testing sessions. ERPs are plotted for three types of target strings: target words that were semantically related to the prime word (solid line), target words that were not semantically related to the prime (small dashes), and pseudoword targets (large dashes). (Adapted from McLaughlin, Osterhout, & Kim, 2004.)

near the end (~140 hr). We asked the French learners to make word/nonword judgments regarding the words and pseudowords. We found that after only 14 hr of L2 instruction, the pseudowords elicited a robustly larger N400 than the words (Figure 2a, right column). This was true even though learners were at chance levels ( $d' = 0$ ) when deciding if the letter strings were actual words in French. Furthermore, the correlation between hours of instruction and the word/nonword N400 difference was very robust ( $r = .72$ ), suggesting that the N400 difference was approximately linearly related to the learner's exposure to the L2. This result

Q3

suggests that the French learners rapidly extracted enough information about French word forms so that their brains could discriminate between actual words and pseudowords, even if the learners themselves could not do so. Effects of word meaning, manifested as smaller N400s to words preceded by a related word than by an unrelated word, were observed after  $\sim 60$  hr of instruction (Figure 2b, right column). After  $\sim 140$  hr of L2 instruction (Figure 2c, right column), the amplitude of the word/nonword and relatedness differences approximated that typically observed with similar stimuli in native English or French speakers, even though the learners' explicit word/nonword judgments remained very poor ( $d' < 1$ ). No differences in the N400 amplitude across conditions were observed for a group of subjects who had received no French instruction (Figure 2a–2c, left column). These results show that the L2 learners extracted information about word form after just  $\sim 14$  hr of instruction and extracted information about word meaning by at least  $\sim 60$  hr. Furthermore, these changes in brain activity immediately started to approximate the responses seen in native speakers to analogous stimuli and occurred even while the learner's conscious lexicality judgments were very poor. However, what, exactly, were these learners learning about words? The very early learning observed in this study might involve elemental aspects of linguistic knowledge (e.g., knowledge about word forms) that serve as prerequisites to L2 competence. People rapidly extract co-occurrence statistics for letter and sound combinations within a language (Saffran, Johnson, Aslin, & Newport, 1999). Perhaps the word/nonword N400 effect reflected a similar type of learning. To test that idea, we computed correlations between the word/nonword N400 effect and bigram and trigram frequencies (as computed from the relevant portions of the French text). These correlations were very weak ( $p > .3$ ). Another possibility is that learners memorized whole-word forms they were exposed to during their L2 instruction. This idea receives some support from the robust correlation between the N400 word/nonword effect and the frequency of the target words in the learners' French text ( $p = .67$ ).

## 212 *Novice Learners, Longitudinal Designs, ERPs*

### ERPs and L2 Morphosyntactic Learning

A second line of research in our laboratory is motivated by the observation, discussed earlier, that syntactic and semantic anomalies embedded within sentences elicit distinct ERP effects (the P600 and N400 effects, respectively). This finding suggests that separable syntactic and semantic processes exist. In an L2 learning context, one implication is that L2 learners must somehow segregate linguistic input into those aspects of the language that relate to sentence form and those that relate to sentence meaning; that is, learners “grammaticalize” some aspects of the L2, but not others. Such a conclusion introduces a host of interesting questions. How and when do L2 learners grammaticalize aspects of the L2? Which aspects of the syntax are grammaticalized first? What factors influence the rate and eventual success of grammaticalization? How similar is the process of grammaticalization in learning an L1 or an L2? In our work, what we mean by “grammaticalization” is specifically the instantiation of grammatical knowledge into the learner’s online, real-time language processing system. Our assumption is that once a feature of the L2 has been grammaticalized, violations of that aspect of the grammar should elicit a P600 effect.

To investigate such questions, we have focused on one particular aspect of syntactic learning, namely the acquisition of grammatical features and their associated morphosyntactic rules. The centrality of grammatical features in recent formal theories has led to renewed interest in studying the acquisition of the grammatical morphemes that encode these features. Within linguistic theory (Radford, 1988), semantic properties of language are encoded in word (lexical) categories such as Noun (N), Verb (V), and Adjective (A); these categories are shared by all languages. Grammatical properties, by contrast, are encoded in the morphology of grammatical categories such as noun phrase (NP) and verb phrase (VP). Grammatical categories are associated with grammatical features such as gender, number, and verbal person. These features (and how they are involved in morphosyntax)

vary across languages (Table 1); for example, with respect to grammatical features, English has a number feature in the NP (e.g., *boy* vs. *boys*), whereas French has both number and gender. With respect to morphosyntax, English and French both have sentential agreement (i.e., agreement of the verb with the subject in verbal person and number; e.g., *I like* vs. *He likes*), but only French has NP agreement (i.e., agreement between the noun and its determiner/adjective in number and gender; e.g., *le garçon* vs. *les garçons*; excluding the restricted English case of *this/those*) (Table 1). Languages also differ with respect to the explicitness of grammatical features; for example, although English and French both show agreement between subject and verb, the explicitness of the grammatical marking differs. English present tense uses only *-s* in third person singular (*he walks, they/I/etc. walk*). French present uses five orthographic forms (*je marche, tu marches* “I/you walk”), but has only three distinct pronunciations. As for NP agreement, English marks number on the noun alone (*the little boy[s]*), whereas French marks number (and gender) on the agreeing determiner/adjective (*le[s] petit[s] garçon[s]*) (Table 1). In oral French, however, plurality can only be heard on the determiner. English has no grammatical gender feature, whereas French does. What factors might inhibit or facilitate grammaticalization of these features and their morphosyntactic rules? One frequent claim is that only features that are present in the L1 can be acquired during L2 acquisition (Franceschina, 2001; Hawkins & Franceschina, 2004). Other researchers, however, argue that novel L2 features can be learned, albeit more slowly than those that are present in the L1 (White, 2003). Thus, there is no consensus about whether, or when during acquisition, L2 learners acquire L2 features that are not present in their L1. Even less is known about the acquisition of L2 morphosyntactic rules (e.g., agreement in number or gender) that are not present in the L1.

Another factor that seems likely to play a role in L2 grammatical morpheme learning is the covariation between morphology and phonology. The interaction between morphology and

214 *Novice Learners, Longitudinal Designs, ERPs*

Table 1

*Sentential verb-subject agreement: regular verbs, present tense*

	English	French
Singular		
1st	I like	Je marche
2nd	You like	Tu marche(s)
3rd	He likes	Il marche
Plural		
1st	We like	Nous marchons
2 <sup>nd</sup>	You like	Vous marchez
3rd	They like	Ils marche(nt)
Noun phrase (NP) number agreement		
Singular	the boy	le petit garçon
Plural	the boys	les petit(s) garçon(s)
Noun phrase (NP) gender agreement		
Male	NA	le garçon
Female	NA	la fille

*Note.* Morphemes in parentheses are not expressed Phonologically.

phonology can be clearly seen in written French. French has an opaque orthography due to many suffixes being phonologically silent (Table 1). Thus, the plural suffix *-s*, which marks the plural across all elements in the NP (*le-s jeune-s fille-s* “the young girls”) is silent on the noun (as well as on the adjective) in almost all instances. A similar situation arises in the VP, where variations in verbal person are marked orthographically on the verb but are silent in most oral forms. Thus, the different inflections for a regular verb such as *marcher* (to walk) sound identical across three different persons/spellings (Table 1). The effect of the “missing” phonological cue is notorious on spelling. French children are drilled in the morphological variations throughout elementary school, and yet they continue to make grammatical errors, by failing to add the plural inflection in writing. Errors such as *les chien* or *Ils mange* are frequent (Negro & Chanquoy,



2000) and can be seen in adults as well. Errors are much rarer when phonology is available as a cue (Largy & Fayol, 2001); for example, in the VP, confusions are not made among persons for *nous* and *vous* (“we/you” formal or plural), as these forms are both morphologically and phonologically distinct from each other and all other persons. Errors are also far less frequent when the morphological ending is paired with a phonological difference, as is the case for certain irregular and stem-changing verbs in French. The results from French are in line with results from studies of Dutch spelling (Frisson & Sandra, 2002; Sandra, Frisson, & Daems, 1999) showing that adult Dutch writers continue to confuse the first and third person singular for verbs where these two persons have distinct morphological endings but identical oral realizations (e.g., *rijd* “drive” and *rijdt* “drives,” which are both pronounced as/rɛIt/), but they do not confuse the two for verbs where they have distinct phonologies (e.g., *werkt* /werkt/“works” and *werk* /werk/“works”). In adult L2 acquisition, the use of oral cues has been advocated by linguists to enhance the learning of certain morphological rules, such as verb and adjective variations (Herschensohn, 1993). Arteaga, Gess, and Herschensohn (2003) found that English learners of French who were instructed to note phonological variations of masculine and feminine forms of variable adjectives in French were quicker to learn the morphological rule than were students who were not given such instruction but were taught only the spelling alterations between the two forms.

Therefore, one reasonable prediction is that L2 learners will acquire an L2 feature or morphosyntactic rule more quickly when the relevant inflectional morphology is phonologically realized. However, this possibility has received little direct attention in the recent L2 literature. It also seems likely that L1-L2 similarity and phonological-morphological covariation might have interactive effects during L2 learning; for example, L1-L2 similarity combined with phonological realization of the relevant grammatical morphemes might lead to very fast learning, whereas L1-L2 dissimilarity combined with no phonological realization might lead to very slow learning.

216 *Novice Learners, Longitudinal Designs, ERPs*

Finally, various theorists have suggested that L2 morphosyntactic learning involves discontinuous stages (Myles, Hooper, & Mitchell, 1998; Vainikka & Young-Scholten, 1996; Wong-Fillmore, 1976; Wray, 2001). This seems to be true of children learning their L1, who begin by memorizing particular combinations of words and only later induce general syntactic rules (Tomasello, 2000). With respect to L2 learning, one account holds that morphologically complex words are initially learned by rote and memorized as unanalyzed chunks (Myles et al., 1998; Wray, 2002). Eventually, learners unpack these chunks into roots and grammatical morphemes, induce grammatical rules governing their use, and deductively use these units to produce novel utterances. To be specific, an L2 learner might initially memorize the fact that certain subjects are followed by certain forms of the verb, without decomposing the verb into root + inflection or applying a general morphosyntactic agreement rule. In this stage of learning, the learner associates meanings with the undecomposed word and either memorizes the two words as a chunk or learns about word sequence probabilities (e.g., that *Tu* is followed by *marches*, whereas *Il*s is followed by *marchent*).<sup>3</sup> Unfortunately, when using conventional behavioral methods, it is very difficult to identify the transition from rote-memorization to rule use (Wray, 2001). Fortunately, ERPs might be an ideal tool for testing this theoretical claim of developmental discontinuity. As noted earlier, the N400 amplitude is highly sensitive to novel words and to word sequence probabilities, whereas the P600 effect is sensitive to grammatical rule violation. If an L2 learner initially memorizes salient word sequences, then unfamiliar word combinations (e.g., *\*Tu adorez*) should produce larger N400s than familiar ones (e.g., *Tu adores*). If the learner eventually decomposes the verb into root + morpheme and induces the rule, use of the wrong verbal form should elicit a P600 effect. In essence, the transition for N400 to P600 would reflect the transition from rote-memorized (i.e., *lexicalized*) knowledge to *grammaticalized* knowledge.

We investigated these predictions using a longitudinal experimental design involving 14 English-speaking novice French

Q4

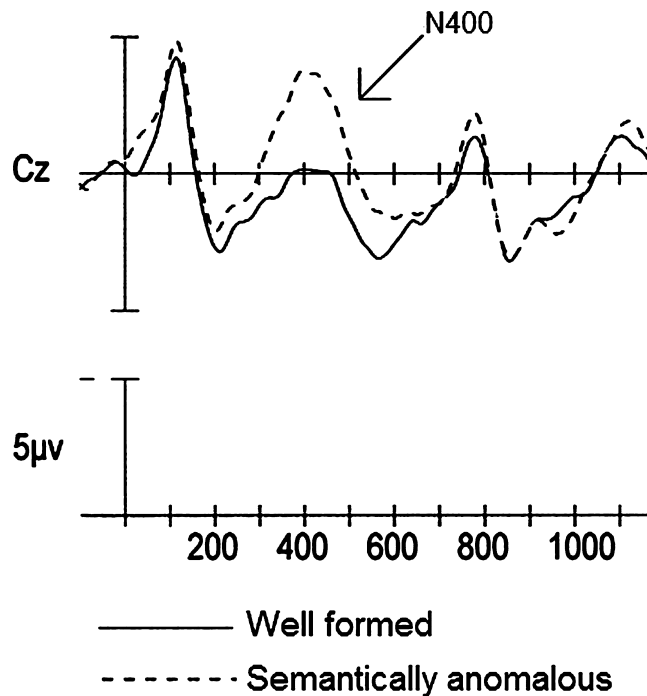
Q5

learners. Each learner was tested after approximately 1 month, 4 months, and 8 months of university classroom French instruction. Our stimuli were as follows:

- (1) *Sept plus cinq\?livre font douze.*  
semantic condition
- (2) *Tu adores\\*aderez le français.*  
verbal person agreement condition/phonologically realized
- (3) *Tu manges des hamburgers\\*hamburger pour dîner.*  
number agreement condition/phonologically unrealized

In (1), the noun *livre* is semantically anomalous. In (2), the verb *aderez* is conjugated incorrectly, given the preceding sentence fragment. In (3), the noun *hamburger* disagrees with the syntactic number of the plural article. Our stimuli were selected from the material in the textbook assigned during the first month of instruction. The anomalous items in the verbal person condition involved a grammatical rule that was present in the L1 and an orally realized contrast between inflectional morphemes. The anomalous items in the number agreement condition involved a rule that was not present in the L1 and a phonologically unrealized contrast between inflectional morphemes. Therefore, our prediction was that L2 learners would respond to the anomaly in (2) with less L2 exposure compared to the anomaly in (3).

As expected, the native-French speakers showed an N400 effect to the semantically anomalous words and large P600 effects to the two types of syntactic anomaly. The learners, as is often the case, showed striking individual differences, both in a behavioral “sentence acceptability judgment” task and in the pattern of ERPs elicited by the anomalous stimuli. We segregated the learners into upper (“fast learners”) and lower (“slow learners”) halves, based on their performance in the sentence-acceptability judgment task that occurred concurrently with ERP data collection, and averaged the ERPs separately for each group. Results for the “fast learner” group will be described here. At each testing session,

218 *Novice Learners, Longitudinal Designs, ERPs*

*Figure 3.* ERPs to critical words in the well-formed (solid line) and semantically anomalous (dashed line) conditions, collapsed over the three testing sessions.

including the initial session that occurred after just 1 month of instruction, semantically anomalous words elicited a robust N400 effect, and this effect changed minimally with increasing instruction (Figure 3, averaged over the three testing sessions). Results for the verbal person condition are shown in Figure 4. After just 1 month of instruction, the learners' brains discriminated between the syntactically well-formed and ill-formed sentences. However, rather than eliciting the P600 effect (as we saw in native-French speakers), the syntactically anomalous words elicited an N400-like effect. (This effect did not differ in distribution from the N400 effect elicited by the semantically anomalous words.) By 4 months, the N400 effect was replaced by a

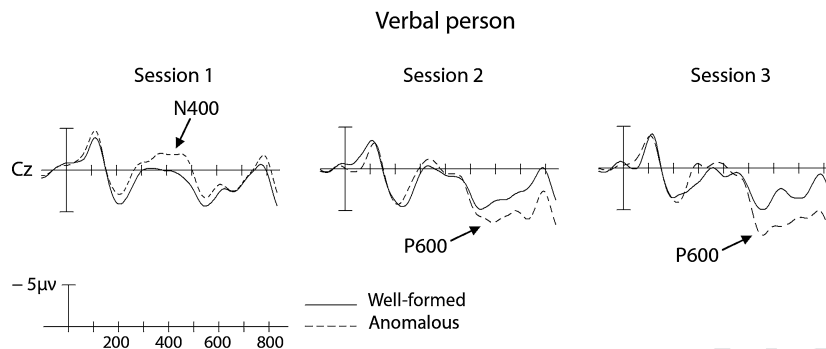


Figure 4. ERPs to critical words in the well-formed (solid line) and verbal person anomaly (dashed line) conditions, plotted separately for each of the three testing sessions.

P600-like positivity. Results for the number agreement condition can be summarized easily: Learners performed very poorly in the sentence acceptability judgment task for these materials, and there were no robust differences between the agreeing and disagreeing stimuli (Figure 5, averaged over three testing sessions).

These results are consistent with the predictions that we were testing. First, L1-L2 similarity combined with phonological realization of the relevant grammatical morphemes produced very fast L2 syntactic learning, whereas L1-L2 dissimilarity combined with no phonological realization produced very slow learning. This occurred even though our learners were drilled repeatedly on both rules from nearly the first day in class. However, the two rules we tested represent the ends of a putative continuum of morphosyntactic difficulty; without additional data, it is impossible to know whether L1-L2 similarity or phonological realization of grammatical morphemes had a larger impact on the learning rate. What can be said with more certainty is that at least some L2 syntactic rules are learned amazingly quickly. Apparently, at least some types of L2 rule are incorporated into the online sentence processing system after a very small amount of L2 instruction.

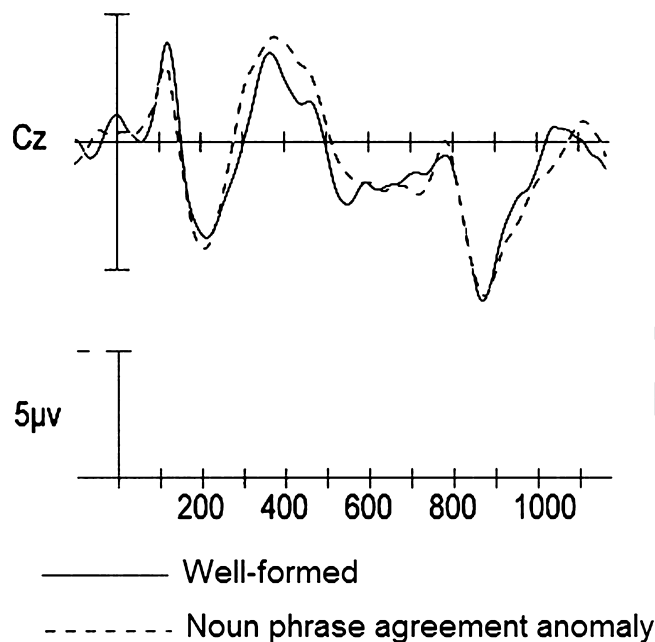
220 *Novice Learners, Longitudinal Designs, ERPs*

Figure 5. ERPs to critical words in the well-formed (solid line) and noun phrase agreement anomaly (dashed line) conditions, collapsing over the three testing sessions.

The second prediction was that L2 morphosyntactic anomalies would show a discontinuous pattern over time: Early in learning, such anomalies would elicit an N400 effect in learners, whereas later in learning, these same anomalies would elicit a P600 effect. This pattern is exactly what we observed for the verbal person anomalies. If our interpretation is correct, then our adult L2 learners grammaticalized this aspect of the L2 after just a few months of L2 instruction. Our results can be explained by assuming that learners (much like child L1 learners) initially memorize salient word sequences (e.g., *tu adores*). Violations of the verbal person rule (e.g., *tu adorez*) result in novel word combinations and, hence, elicit an N400 effect. After more instruction, learners induce a general verbal person rule (*tu -s, nous -ons, vous -ez*, etc.); violations of the rule elicit a P600 effect. The

P600 effect occurred at similar temporal latencies for the natives and learners. Thus, after just 80 hr of instruction, learners' and natives' ERP responses to verbal person violations were qualitatively similar.

English and French are highly similar in many respects; for example, they share many cognates and a similar (although not identical) system of morphosyntax, and they both belong to the Indo-European family. English and Finnish, by contrast, share very few cognates and come from different language families. English has a relatively impoverished morphosyntax, whereas Finnish has an extensive system: Nouns are marked for number, case, and possessiveness, and finite verbs are marked for tense or mood, number, and person. Agreement rules in Finnish exist for number, person, and case. Finnish is generally thought to have a subject-verb-object canonical structure. However, because the morphosyntactic system usually makes the syntactic and semantic functions of each constituent in a clause unambiguous, Finnish speakers have considerable flexibility in ordering words. Noncanonical structures are permitted to mark a change in focus or for other purposes. Furthermore, the Finnish morphosyntactic system is highly regular and fully realized in the phonology. Inflectional suffixes are regular, and they are added to the word stems in a fixed order. However, the addition of the endings is sometimes accompanied by sound alternations in the stem of the word.

These properties of Finnish morphosyntax provide a useful contrast with French, in the context of a native English-speaking L2 learner. If the L1-L2 similarity is the primary determinant of the rate of L2 syntactic learning, then learners should be quite slow to acquire the many novel aspects of Finnish morphosyntax. If, conversely, regularity, transparency, and phonological realization in the morphosyntax are more important, then learners might acquire the system rather quickly.

Recent work has shown that morphosyntactic and semantic anomalies elicit the P600 and N400 effects, respectively, when the subjects are native speakers of Finnish (Palolahti, Leino, Jokela,

222 *Novice Learners, Longitudinal Designs, ERPs*

Kopra, & Paavilainen, 2005). In our lab, we have begun to study English-speaking university students as they progress through their first year of Finnish instruction. We hope to determine how much Finnish instruction is needed before morphosyntactic violations elicit an ERP linguistic anomaly effect. We are presenting sentences that contain either a semantic anomaly and one of two kinds of morphosyntactic error, as illustrated below:

(4a) Sinä (2nd p. sing.) syöt (2nd p. sing.) lihaa.  
“You eat meat.”

(4b) \*Sinä (2nd p. sing.) syön (1st p. sing.) lihaa.  
\*“You eat meat.”

(5a) He istuvat kahvilassa (inessive case) illalla.  
“They sit *in* a café at night.”

(5b) \*He istuvat kahvilasta (elative case) illalla.  
\*“They sit *from* a café at night.”

In (4b), the verb is in the wrong person form. In (5b), the noun is attached to an incorrect case particle. Because English has a verbal person rule but not a system of explicit case marking, learning might be relatively fast for Finnish verbal person but very slow for case if L1-L2 similarity is the primary determinant of learning rate. Conversely, if phonological realization of the grammatical morphemes is very important, then both the similar and dissimilar aspects of the morphosyntactic system might be learned relatively quickly. In a pilot study, we tested five English speakers who had just completed their initial 9 months of Finnish instruction. The results were as follows: Semantic anomalies elicited an N400 effect; verbal person anomalies elicited a robust P600 effect; and case anomalies elicited a P600-like effect, although with a much smaller amplitude. These preliminary data suggest that even unfamiliar aspects of a complex morphosyntactic system can be grammaticalized within the first year of classroom instruction. Our ongoing research with this population might tell us more about the rate of learning and



whether this type of morphosyntactic learning is continuous or discontinuous.

### Conclusions and Final Comments

Our goal in this article has been to describe a research paradigm for investigating what L2 learners learn as they progress through a period of L2 instruction and when they learn it. By learning, we mean specifically the incorporation of L2 knowledge into the learner's online, real-time language processing system. We do not claim to have definitive answers to any of the questions that we hope to answer. For example, our pilot Finnish study suggests that English-speaking L2 learners might "grammaticalize" novel aspects of Finnish morphosyntax during their initial 8 months of classroom instruction. We have suggested that this might be due, in part, to the fact that the relevant grammatical morphemes are reliably (and uniquely) expressed in the language's phonology. However, of course, there are many other explanations for this effect, if it turns out to be robust. Our proposed paradigm is limited in other ways as well. Although the paradigm minimizes between-subjects variability, it cannot eliminate it. As we found in our French syntax learning study, individual differences continue to exert an influence on learning even when we longitudinally study novice learners in a homogeneous classroom environment. This is not unexpected. Our hope is that our design reduces between-learner variability to a manageable level; eventually, careful study might reveal the primary sources of the remaining variability. Finally, it is important to explicitly recognize that ERPs (like all methods of investigation) imperfectly reflect the cognitive and neural processes underlying L2 learning and processing.

Even given these caveats, our preliminary findings are promising. Our findings suggest that it is possible to ascertain *when* some aspect of the L2 has been incorporated into the set of processes that allow the learner to comprehend the language

224 *Novice Learners, Longitudinal Designs, ERPs*

in real time. Moreover, our results suggest that developmental discontinuities exist with respect to *what* has been learned: early in the learning process, learners seem to memorize frequently occurring words or groups of morphemes. With a little more L2 exposure, learners associate the memorized lexical forms with meanings, begin to decompose the structurally complex units into roots plus grammatical morphemes, and learn the rules for agreement among these elements. This type of learning can occur with remarkably little L2 exposure, in a standard classroom setting that is often thought to provide a less-than-ideal learning environment. Some learners even seem to quickly grammaticalize novel aspects of a complex L2 morphosyntactic system that is unlike anything in their L1. Our results are preliminary, but they clearly conflict with the conventional belief that adult L2 syntactic learning is generally slow and problematic, especially for aspects of the L2 that are not present in the L1. The conventional belief about the difficulty of L2 syntactic learning is based, in part, on L2 production: Many speakers do have difficulty with grammatical aspects of an L2 when they try to speak it. However, one general rule about L1 learning is that a learner's ability to understand the language develops in advance of her ability to produce it. It seems likely that this maxim also applies to L2 learners. If so, then perhaps there is no contradiction between our data (indicating a rapid acquisition and implementation of some types of L2 knowledge during language comprehension) and the common impression that L2 learners have a difficult time producing certain aspects of the L2.

We should note that other researchers have come part way in implementing our proposed research paradigm, although (to the best of our knowledge) no one has implemented it completely. Several groups of researchers have used ERPs to study the incorporation of L2 knowledge into the online language processing system (e.g., Hahne, 2001). However, the learners in these studies were not novice learners, but had extensive (and possibly highly variable) experience with their L2. Furthermore, none of these studies provided a contrast between different levels of L2 exposure or proficiency. In other studies, researchers have

recorded ERPs from subjects who were acquiring aspects of an artificial language (Friederici, Steinhauer, & Pfeifer, 2002; McCandliss, Posner, & Givon, 1997). As in our studies described earlier, the learners were novices with respect to the “language” to be learned. Although these studies are quite interesting, it is unclear to what degree the results can be generalized to the acquisition of a natural language in a typical (within the United States) L2 learning environment. These studies have also not included a contrast involving learners at different stages of acquisition.

Only a few studies have both recorded ERPs and contrasted different stages of L2 exposure or proficiency. For example, Weber-Fox and Neville (1996) examined the ERP response to linguistic anomalies in adult Chinese-English bilinguals who had been exposed to English at various ages, ranging from 1 to 16 years. They report interesting differences in the ERP responses to a variety of linguistic anomalies. However, because the study was cross-sectional in nature, there is inevitably some ambiguity about whether these different ERP patterns are in fact manifestations of age of exposure effects or, instead, reflect some other source of uncontrolled subject variability.

The approach proposed here is not a panacea and is itself open to criticism. For example, at the moment we cannot predict or explain learner differences of the type we observed in our French syntax learning study. It is also true that longitudinal designs are extremely time-consuming and are not always viable. Nonetheless, this research paradigm might produce novel results that are less subject to the ambiguity plaguing much of the existing literature. By focusing on changes in the neural manifestations of L2 processing that occur with increasing L2 instruction within a set of L2 learners, we hope to demonstrate the value of asking the what and when questions: Exactly what processes change over time with increasing exposure to or proficiency in an L2? When (in the course of understanding a word or sentence and in the course of learning a language) do those changes occur? Some progress has already been made toward answering such questions in the domain of L2 speech perception (Kuhl, 2004; Osterhout

226 *Novice Learners, Longitudinal Designs, ERPs*

et al., 2004). Perhaps we can reasonably hope that similar advances will soon be forthcoming with respect to the online processing of L2 words and sentences.

## Notes

<sup>1</sup>Friederici and colleagues (Hahne & Friederici, 1999; Friederici, 1995) have proposed a two-stage functional model for these ERP effects. They claim that the anterior negativity reflects a fast, automatic syntactic analyzer, and the P600 reflects syntactic reanalysis. As we have explained elsewhere (Osterhout et al., 2004, pp. 294–298), we do not believe that the available evidence strongly supports these theoretical claims.

<sup>2</sup>Another problematic aspect of using the LAN effect to study L2 processing is that there are a significant number of reports involving native speakers in which they are not reported (e.g., Ainsworth-Darnell, Shulman, & Boland, 1998; Allen et al., 2003; Hagoort, Brown, & Groothusen, 1993; Kim & Osterhout, 2005; Kuperberg et al., 2003; McKinnon & Osterhout, 1996; Osterhout, Bersick, & McLaughlin, 1997; Osterhout & Mobley, 1995). This makes it difficult to interpret the absence of the effect in L2 learners.

<sup>3</sup>A related account of L2 learning (Vainikka & Young-Scholten, 1996) proposed that L2 learners initially transfer their L1 lexical categories, but not their L1 grammatical features and categories; learners are claimed to infer these features and categories from the presence of the associated L2 grammatical morphemes. Thus, syntactic structure is projected from the acquired lexical material, and learners must go through a stage of lexical learning before they acquire the features and associated morphosyntax (see also Clahsen, Penke, & Parodi, 1994).

## References

- Ainsworth-Darnell, K., Shulman, R., & Boland, J. (1998). Dissociating brain responses to syntactic and semantic anomalies: Evidence from event-related brain potentials. *Journal of Memory and Language*, 38, 112–130.
- Allen, M. D., Badecker, W., & Osterhout, L. (2003). Morphological analysis during sentence processing. *Language and Cognitive Processes*, 18, 405–430.
- Arteaga, D., Gess, R., & Herschensohn, J. (2003). Focusing on phonology to teach morphological form in French. *Modern Language Journal*, 8, 58–70.
- Bentin, S. (1987). Event-related potentials, semantic processes, and expectancy factors in word recognition. *Brain and Language*, 31, 308–327.

- Bentin, S., McCarthy, G., & Wood, C. (1985). Event-related potentials, lexical decision, and semantic priming. *Electroencephalography and Clinical Neurophysiology*, *60*, 343–355.
- Chee, M. W. L., Tan, E. W. L., & Thiel, T. (1999). Mandarin and English single word processing studies with functional magnetic resonance imaging. *Journal of Neuroscience*, *19*, 3050–3056.
- Clahsen, H., Penke, M., & Parodi, T. (1994). Functional categories in early child German. *Language Acquisition*, *3*, 395–429.
- Franceschina, F. (2001). Morphological or syntactic deficits in near-native speakers? An assessment of some current proposals. *Second Language Research*, *17*, 213–247.
- Friederici, A. D. (1995). The time course of syntactic activation during language processing: A model based on neuropsychological and neurophysiological data. *Brain and Language*, *50*, 259–284.
- Friederici, A. D., Steinhauer, K., & Pfeifer, E. (2002). Brain signatures of artificial languages: Evidence challenging the critical period hypothesis. *Proceedings of the National Academy of Sciences of the United States of America*, *99*, 529–534.
- Frisson, S., & Sandra, D. (2002). Homophonic forms of regularly inflected verbs have their own orthographic representations: A developmental perspective on spelling errors. *Brain and Language*, *81*, 545–554.
- Grosjean, F. (1998). Studying bilinguals: Methodological and conceptual issues. *Bilingualism: Language and Cognition*, *1*, 131–149.
- Hagoort, P., Brown, C., & Groothusen, J. (1993). The syntactic positive shift (SPS) as an ERP measure of syntactic processing. *Language and Cognitive Processes*, *8*, 439–483.
- Hahne, A. (2001). What's different in second-language processing? Evidence from event-related brain potentials. *Journal of Psycholinguistic Research*, *30*, 251–266.
- Hahne, A., & 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.
- Hahne, A., & Friederici, A. D. (2001). Processing a second language: Late learners comprehension mechanisms as revealed by event-related brain potentials. *Bilingualism: Language and Cognition*, *4*, 123–141.
- Hawkins, R., & Franceschina, F. (2004). Explaining the acquisition and non-acquisition of determiner-noun gender concord in French and Spanish. In P. Prevost & J. Paradis (Eds.), *The acquisition of French in different contexts* (pp. 175–206). Amsterdam: Benjamins.
- Herschensohn, J. (1993). Applying linguistics to teach morphology: Verb and adjective inflection in French. *International Review of Applied Linguistics*, *30*, 97–112.

228 *Novice Learners, Longitudinal Designs, ERPs*

- Kim, A., & Osterhout, L. (2005). The independence of combinatory semantic processing: Evidence from event-related potentials. *Journal of Memory and Language*, *52*, 205–225.
- Kim, K. H. S., Relkin, N. R., Lee, K.-M., & Hirsch, J. (1997). Distinct cortical areas associated with native and second languages. *Nature*, *388*, 171–174.
- Klein, D., Milner, B., Zatorre, R., Zhao, V., & Nikelski, J. (1999). Cerebral organization in bilinguals: Chinese-English verb generation. *NeuroReport*, *10*, 2841–2846.
- Kroll, J., & Sunderman, G. (2003). Cognitive processes in second language learners and bilinguals. The development of lexical and conceptual representations. In C. Doughty & M. Long (Eds.), *The handbook of second-language acquisition* (pp. 104–129). Malden, MA: Blackwell.
- Kuhl, P. (2004). Early language acquisition: Cracking the speech code. *Nature Reviews Neuroscience*, *5*, 831–843.
- Kuperberg, G., Holcomb, P. J., Sitnikova, T., Greve, D., Dale, A. M., & Caplan, D. (2003). Distinct patterns of neural modulation during the processing of conceptual and syntactic anomalies. *Journal of Cognitive Neuroscience*, *15*, 272–293.
- Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic anomaly. *Science*, *207*, 203–205.
- Kutas, M., & Hillyard, S. A. (1984). Brain potentials during reading reflect word expectancy and semantic association. *Nature*, *307*, 161–163.
- Largy, P., & Fayol, M. (2001). Oral cues improve subject-verb agreement in written French. *International Journal of Psychology*, *36*, 121–131.
- McCandliss, B. D., Posner, M. I., & Givon, T. (1997). Brain plasticity in learning visual words. *Cognitive Psychology*, *33*, 88–110.
- McKinnon, R., Allen, M., & Osterhout, L. (2003). Morphological decomposition involving nonproductive morphemes: ERP evidence. *NeuroReport*, *14*, 883–886.
- McKinnon, R., & Osterhout, L. (1996). Constraints on movement phenomena in sentence processing: Evidence from event-related brain potentials. *Language and Cognitive Processes*, *11*, 495–523.
- McLaughlin, B. (1998). Second language learning revisited: The psycholinguistic perspective. In A. F. Healy & L. E. Bourne, Jr. (Eds.), *Foreign language learning: Psycholinguistic studies on training and retention* (pp. 399–412). Mahwah, NJ: Erlbaum.
- McLaughlin, J., Osterhout, L., & Kim, A. (2004). Neural correlates of second-language word learning: Minimal instruction produces rapid change. *Nature Neuroscience*, *7*, 703–704.

- Myles, F., Hooper, J., & Mitchell, R. (1998). Rote or rule? Exploring the role of formulaic language in classroom foreign language learning. *Language Learning, 48*, 323–363.
- Negro, I., & Chanquoy, L. (2000). Subject-verb agreement with present and imperfect tenses: a developmental study from 2<sup>nd</sup> to 7<sup>th</sup> grade. *European Journal of Psychology of Education, 15*, 113–134.
- Neville, H. J., Nicol, J. L., Barss, A., Forster, K. I., & Garrett, M. (1991). Syntactically based sentence processing classes: Evidence from event-related brain potentials. *Journal of Cognitive Neuroscience, 3*, 151–165.
- Osterhout, L., Bersick, M., & McLaughlin, J. (1997). Brain potentials reflect violations of gender stereotypes. *Memory and Cognition, 25*, 273–285.
- Osterhout, L., & Holcomb, P. J. (1992). Event-related brain potentials elicited by syntactic anomaly. *Journal of Memory and Language, 31*, 785–806.
- Osterhout, L., & Holcomb, P. J. (1993). Event-related potentials and syntactic anomaly: Evidence of anomaly detection during the perception of continuous speech. *Language and Cognitive Processes, 8*, 413–438.
- Osterhout, L., McLaughlin, J., Kim, A., Greenwald, R., & Inoue, K. (2004). Sentences in the brain: Event-related potentials as real-time reflections of sentence comprehension and language learning. In M. Carreiras, & C. Clifton, Jr. (Eds.), *The on-line study of sentence comprehension* (pp. 271–308). New York: Psychology Press.
- Osterhout, L., & Mobley, L. A. (1995). Event-related brain potentials elicited by failure to agree. *Journal of Memory and Language, 34*, 739–773.
- Osterhout, L., & Nicol, J. (1999). On the distinctiveness, independence, and time course of the brain responses to syntactic and semantic anomalies. *Language and Cognitive Processes, 14*, 283–317.
- Palolahti, M., Leino, S., Jokela, M., Kopra, M., & Paavilainen, P. (2005). Event-related potentials suggest early interaction between syntax and semantics during on-line sentence comprehension. *Neuroscience Letters, 384*, 222–227.
- Perani, D., Paulesu, E., Galles, N. S., Dupoux, E., Dehaene, S., Bettinardi, V., et al. (1998). The bilingual brain: Proficiency and age of acquisition of the second language. *Brain, 121*, 1841–1852.
- Potter, M. C., So, K.-F., Von Eckardt, B., & Feldman, L. B. (1984). Lexical and conceptual representation in beginning and more proficient bilinguals. *Journal of Verbal Learning and Verbal Behavior, 23*, 23–38.
- Raboyeau, G., Marie, N., Baldyck, S., Gros, H., Demonet, J., & Cardebat, D. (2004). Lexical learning of the English language: A PET study in healthy French subjects. *NeuroImage, 22*, 1808–1818.
- Radford, A. (1988). *Transformational grammar*. Cambridge: Cambridge University Press.

230 *Novice Learners, Longitudinal Designs, ERPs*

- Saffran, J. R., Johnson, E. K., Aslin, R. N., & Newport, E. L. (1999). Statistical learning of tonal sequences by human infants and adults. *Cognition*, 70, 27–52.
- Sandra, D., Frisson, S., & Daems, F. (1999). Why simple verb forms can be so difficult to spell: The influence of homophone frequency and distance in Dutch. *Brain and Language*, 68, 277–283.
- Tomasello, M. (2000). *Trends in Cognitive Sciences*, 4, 156–160. Q7
- Vainikka, A., & Young-Scholten, M. (1996). Gradual development of L2 phrase structure. *Second Language Research*, 12, 7–39.
- Weber-Fox, C., & Neville, H. J. (1996). Maturation constraints on functional specializations for language processing: ERP and behavioral evidence in bilingual speakers. *Journal of Cognitive Neuroscience*, 8, 231–257.
- White, L. (2003). *Second language acquisition and Universal Grammar*. Cambridge: Cambridge University Press.
- Wong-Fillmore, L. (1976). *The second time around: Cognitive and social strategies in second language acquisition*. Unpublished doctoral dissertation, Stanford University, Stanford, CA.
- Wray, A. (2002). *Formulaic language and the lexicon*. Cambridge: Cambridge University Press.



## Queries

- Q1** Author: Check last name.
- Q2** Author: Provide university departments for all of the authors.
- Q3** Author: (a), (b), ... do not appear in art. Please insert.
- Q4** Author: Wray, 2002 in ref. List. Please check throughout.
- Q5** Author: Necessary to explain use of asterisk?
- Q6** Author: Please cite Hahne & Friederci 2001 in text.
- Q7** Author: Provide article title for Tomasello article.