An ERP Study of Second Language Learning after Childhood: Effects of Proficiency

Shiro Ojima¹,², Hiroki Nakata¹, and Ryusuke Kakigi¹

Abstract

Whether there is an absolute critical period for acquiring language is a matter of continuous debate. One approach to address this issue is to compare the processes of second language (L2) learning after childhood and those of first language (L1) learning during childhood. To study the cortical process of postchildhood L2 learning, we compared event-related brain potentials recorded from two groups of adult Japanese speakers who attained either high or intermediate proficiency in English after childhood (J-High and J-Low), and adult native English speakers (ENG). Semantic anomalies embedded in English sentences evoked a clear N400 component in all three groups, with only the time course of the brain activation varying among the groups. Syntactic violations elicited a left-lateralized negativity similar to the left anterior negativity in ENG and J-High, but not in J-Low. In ENG, a P600 component was additionally found. These results suggest that semantic processing is robust from early on in L2 learning, whereas the development of syntactic processing is more dependent on proficiency as evidenced by the lack of the left-lateralized negativity in J-Low. Because early maturation and stability of semantic processing as opposed to syntactic processing are also a feature of L1 processing, postchildhood L2 learning may be governed by the same brain properties as those which govern childhood L1 learning. We argue that these processes are qualitatively similar in many respects, with only restricted domains of language processing being subject to absolute critical period effects.

INTRODUCTION

It is still controversial whether there is a “critical period” for language acquisition, after which complete mastery of a language is impossible (Lenneberg, 1967). The offset of this hypothesized time window for language learning was argued to be around puberty, and it is claimed that the knowledge of a second language (L2) attained after childhood is qualitatively different from that of a first language (L1) attained during childhood (Bley-Vroman, 1989). The well-observed tendency for L2 speakers not to achieve native-like mastery of the target language fits well with this prediction. In both behavioral and brain sciences, however, counterarguments have also been provided against the critical period hypothesis. For instance, the possibility that the same principles govern the development of L1 competence and that of L2 competence is explored by theoretical linguists (Epstein, Flynn, & Martohardjono, 1996), who argue that highly proficient L2 speakers differ from native speakers only quantitatively, but not qualitatively. Animal studies also indicate that experience in adulthood can cause substantial alterations even in primary sensory areas, which were previously thought to become most stable after a hypothesized critical period (Gilbert, 1998). Hence, the issue of whether age factors have absolute effects on cortical learning and processing of an L2, as predicted by the critical period hypothesis, requires further inquiry.

To characterize the development of cortical processing of an L2, we employed the event-related brain potentials (ERPs) technique that has high enough temporal resolution to disentangle the distinct substages of language processing. ERPs are records of small voltage changes that reflect neural activities triggered by an external event, and several ERP indices of language processing have been found. For example, words that are incongruous with the preceding semantic context elicit an N400 component (Kutas & Hillyard, 1980). Violations of syntactic constraints evoke one or more of the following components: an early left anterior negativity (ELAN; Hahne & Friederici, 1999), a left anterior negativity (LAN; Münte, Heinze, & Mangun, 1993), and a P600 (Osterhout & Holcomb, 1992). Several ERP studies investigated cortical processing of L2s. Since Ardal, Donald, Meuter, Muldrew, and Luce (1990), an N400 elicited by semantic violations has been consistently found in the ERPs of nonnative speakers. The differences between native and nonnative speakers in semantic processing appeared in most cases as latency differences (Weber-Fox & Neville, 1996, 2001), but never

¹National Institute for Physiological Sciences, Okazaki, Japan, ²University of Essex, Colchester, UK

© 2005 Massachusetts Institute of Technology

Journal of Cognitive Neuroscience 17:8, pp. 1212–1228
as the presence/absence of the N400. The lack of a component evoked by syntactic violations (e.g., LAN, P600), however, is frequently observed in late bilinguals (Hahne, 2001; Hahne & Friederici, 2001; Weber-Fox & Neville, 1996). Furthermore, Weber-Fox and Neville (1996, 2001) suggested that syntactic aspects of sentence processing are more severely affected by the age of immersion in the L2 than semantic aspects.

Positron emission tomography and functional magnetic resonance imaging studies have addressed the effects of L2 proficiency as well as those of age of first exposure. The cortical representation of an L2 in low-proficiency learners is reported to be less lateralized to the left hemisphere compared with that of an L1 (Dehaene et al., 1997), and age of first exposure impacts the frontal regions including Broca’s area more strongly than the temporal areas (Kim, Relkin, Lee, & Hirsch, 1997). However, it is also a common finding that the L1 and the L2 activate overlapping areas of the brain (Vingerhoets et al., 2003), and Perani et al. (1998) claim that the differences in cortical representation previously found between L1 and L2 do not exist in high-proficiency L2 speakers regardless of age of first exposure. Proficiency and age of first exposure are also matters of debate in behavioral studies; early studies that stressed age factors (Johnson & Newport, 1991) have been challenged by later studies claiming that what counts is proficiency (White & Genesee, 1996).

We recorded ERP responses to visually presented English words embedded in sentences, from two groups of nonnative speakers of English who started learning English in Japan at around age 12 but attained either high or intermediate English proficiency (henceforth, J-High and J-Low). A comparison of these two groups enabled us to observe previously unknown effects of L2 proficiency on ERPs that are not conflated with effects of age of first exposure. By viewing J-Low as learners at an intermediate developmental stage and J-High as those at a highly advanced stage, we also inferred developmental patterns in late L2 acquisition. These developmental patterns were then compared with those in child L1 acquisition that were reported elsewhere. ERPs were also recorded from native English speakers (ENG), and we investigated whether late learners could become native-like in their brain responses. Similar developmental patterns across L1 and L2 and attainment of native-like processing in L2, despite the late onset of L2 learning, will argue against the critical period hypothesis. A half of the English sentences, which served as stimuli, contained either a semantic violation (“This house has ten cities in total.”) or a syntactic violation (subject–verb or “S–V” agreement errors: “Turtles moves slowly.”). Inclusion of both types of violations made it possible to explore the possibility that L2 proficiency has different effects on semantic and syntactic processing. ERP responses to words that introduced violations and those to the corresponding control words were selectively averaged, and the effects of violations and group (ENG vs. J-High vs. J-Low) on ERPs were statistically analyzed. We hypothesized that intact semantic and syntactic (agreement) processing, when disrupted by violations, would elicit particular ERP markers of language processing from around the peak of the P2 component onward, namely, an N400 and a complex of LAN–P600, respectively.

RESULTS

Behavior

Table 1 shows the behavioral data (accuracy in the true-or-false [T/F] questions and in the off-line acceptability judgments in the semantic and syntactic conditions) and the results of analyses of variance (ANOVAs) performed on them. The group main effect was significant in all ANOVAs. Planned comparisons among the groups revealed that ENG and J-High, who did not differ from each other, performed better than J-Low in all three tasks (all ps < .01 except for the ENG vs. J-Low contrast.

![Table 1. Behavioral Data and their Statistical Analyses](image-url)

<table>
<thead>
<tr>
<th></th>
<th>ENG</th>
<th>J-High</th>
<th>J-Low</th>
</tr>
</thead>
<tbody>
<tr>
<td>T/F questions</td>
<td>97.4 (2.8)</td>
<td>95.5 (4.0)</td>
<td>89.7 (5.6)</td>
</tr>
<tr>
<td>Off-line judgments in</td>
<td>96.9 (3.5)</td>
<td>96.8 (4.8)</td>
<td>91.9 (5.5)</td>
</tr>
<tr>
<td>SEM (congruous)</td>
<td>89.3 (12.3)</td>
<td>92.9 (9.0)</td>
<td>74.9 (18.0)</td>
</tr>
<tr>
<td>Off-line judgments in</td>
<td>98.7 (2.0)</td>
<td>98.7 (2.1)</td>
<td>92.9 (6.7)</td>
</tr>
<tr>
<td>SEM (incongruous)</td>
<td>80.7 (25.5)</td>
<td>87.1 (18.8)</td>
<td>66.5 (35.0)</td>
</tr>
<tr>
<td>Off-line judgments in</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SYN (correct)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ANOVAs</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Factor</td>
<td>df</td>
<td>F</td>
<td>p</td>
</tr>
<tr>
<td>T/F questions</td>
<td>Group</td>
<td>2,61</td>
<td>17.06</td>
</tr>
<tr>
<td></td>
<td>Condition</td>
<td>1,58</td>
<td>22.07</td>
</tr>
<tr>
<td></td>
<td>Group</td>
<td>2,58</td>
<td>16.06</td>
</tr>
<tr>
<td></td>
<td>C × G</td>
<td>2,58</td>
<td>4.302</td>
</tr>
<tr>
<td>Off-line judgments in</td>
<td>Condition</td>
<td>1,58</td>
<td>24.47</td>
</tr>
<tr>
<td>SEM</td>
<td>Group</td>
<td>2,58</td>
<td>5.318</td>
</tr>
</tbody>
</table>

M = mean; SD = standard deviation; T/F questions = true-or-false questions; SEM = semantic condition; SYN = syntactic condition; C = Condition (congruous vs. incongruous meaning in the semantic condition, and correct vs. incorrect agreement in the syntactic condition); G = Group.
in the syntactic part of the off-line judgments, which resulted in \( p = .038 \). In addition, the subjects responded less accurately to the anomalous sentences in the off-line judgments in both the semantic and syntactic conditions.

**ERPs in the Semantic Condition**

**ERP Morphology**

Figure 1 displays the grand average waveforms in the semantic condition (congruous vs. incongruous). After two early components (P2 at the frontal and central sites and N200 at T5/6), the ERPs to incongruous words show a large negative deflection (N400) in all three groups, with the difference between congruous and incongruous conditions being larger towards the parietal sites. The peak latency and duration of the N400 seem to vary across the groups. In ENG, the ERPs to incongruent words have a late positive component (LPC) following the N400.

P2 Component

An ANOVA on the peak latency of the P2 at the midline sites found a significant condition main effect \( F(1,61) = 8.613, \ p = .005 \), reflecting the earlier peak for the incongruous words, and this tendency was stronger towards Pz as reflected in a significant Condition \times Electrode interaction \( F(2,122) = 4.968, \ e = .863, \ p = .010 \). These effects probably index the onset of the N400 response to incongruous words. As for the peak amplitude of the P2, a significant Hemisphere \times Electrode \times Group interaction was found for the lateral sites \( F(6,183) = 2.503, \ e = .745, \ p = .039 \). This interaction was due to the Hemisphere \times Electrode interaction being significant only in J-High \( F(3,66) = 4.828, \ e = .743, \ p = .010 \).

N400 Component for Incongruous Words

The significant results of all omnibus ANOVAs to be presented below for the semantic condition are listed in Table 2. The analysis of the peak latency at the midline sites found a significant group effect and a significant Electrode \times Group interaction. Subsequent analyses revealed that the effect of group reached significance only at Pz \( F(2,61) = 10.24, \ p < .001 \), reflecting the latency of ENG being earlier than those of J-High and J-Low at this site \( (ps < .001) \). The effect of electrode was significant only in the two nonnative groups \( \text{J-High, } F(2,44) = 5.724, \ p = .006; \text{J-Low, } F(2,44) = 10.12, \ e = .664, \ p = .002 \); this effect reflected a longer latency towards Pz (Figure 2, inset). Essentially identical results were obtained for the peak latency of the N400 at the lateral sites [group effect at P3/4, \( F(2,61) = 11.14, \ p < .001 \); electrode effect in J-High, \( F(2,44) = 8.580, \ e = .845, \ p = .001 \); electrode effect in J-Low, \( F(2,44) = 18.32, \ p < .001 \). The shift of the peak in the nonnative groups is so large that it is reasonable to consider that the frontal peak and the parietal peak actually are of two different components that partially overlap. This is supported by an inspection of individual data (Figure 2, waveforms). In contrast to the peak latency, the peak-to-peak amplitude of the negativity was not influenced by the factor Group. The peak-to-peak amplitude was larger over the right than the left hemisphere.

Subtraction N400

Subtraction waveforms were computed by subtracting the waveform in the congruous condition from that in the incongruous condition (Figure 3), and the peak of the subtraction N400 was identified. A significant group effect was found for the peak latency at the midline sites; planned comparisons revealed that the three groups differed significantly from each other, confirming that the subtraction N400 peaked earliest in ENG and latest in J-Low, with that of J-High coming in between \( (\text{ENG vs. J-High, } p = .001; \text{ENG vs. J-Low, } p < .001; \text{J-High vs. J-Low, } p = .017) \). The latency analyses for the lateral sites found the group main effect and the Hemisphere \times Group interaction significant. Planned comparisons revealed that in the left hemisphere, all three groups differed from each other \( (\text{ENG vs. J-High, } p = .016; \text{ENG vs. J-Low, } p < .001; \text{J-High vs. J-Low, } p = .033) \), whereas in the right hemisphere, ENG differed from J-High and J-Low \( (ps < .001) \). As for the peak amplitude, no significant results involving group were obtained.

Mean Amplitudes: 350–550 msec

At both midline and lateral sites, incongruous words elicited a larger negativity (N400) than congruous words (condition main effect), and this negativity had a centro-parietal distribution (Condition \times Electrode interaction). Analyses of the lateral sites revealed a significant Hemisphere \times Electrode \times Group interaction additionally. Further tests found that the interaction of Hemisphere \times Electrode was significant only in ENG \( F(5,85) = 12.37, \ e = .324, \ p < .001 \) and J-High \( F(5,110) = 6.02, \ e = .585, \ p = .001 \). The interactions in ENG and J-High were due to a hemispheric asymmetry at the anterior sites \( (F7 (left anterior) – F8 (right) = –2.512 \text{ µV in ENG, and } –1.506 \text{ µV in J-High}) \). J-Low almost lacked this asymmetry \( (F7 – F8 = –0.274 \text{ µV}) \) because of a right anterior negativity.

Mean Amplitudes: 550–850 msec

For the midline sites, the Condition \times Electrode interaction, the Condition \times Group interaction, and the
Figure 1. Grand-average ERP waveforms for semantically congruous (thin lines: “This house has ten rooms in total.”) and incongruent words (thick lines: “This house has ten cities in total.”) at all scalp sites. The waveforms of the three subject groups (ENG, J-High, J-Low) are shown separately. Negative is plotted upwards in this and subsequent figures. LPC = late positive component.
In J-High, there was a significant condition effect at Cz \((t(22) = -3.596, p = .002)\) and Pz \((t(22) = -5.162, p < .001)\), reflecting the late N400 response to incongruous words. For the lateral sites, a significant Condition \(\times\) Electrode \(\times\) Group interaction was found. Subsequent tests found that the Condition \(\times\) Electrode interaction was significant in ENG \(F(5,85) = 5.956, \epsilon = .410, p = .006\) and J-Low \(F(5,110) = 13.42, \epsilon = .441, p < .001\). The interactions found in ENG and J-Low were of different nature, however. The one in ENG reflected an LPC for incongruous words with a parietal maximum, which was significant only at P3/4 \([F(1,17) = 7.754, p = .013]\). The Condition \(\times\) Electrode interaction in J-Low reflected the centro-parietal maximum of the late N400. A highly significant condition main effect was also found in J-Low \([F(1,22) = 19.46, p < .001]\). The omnibus ANOVA also found a significant Condition \(\times\) Hemisphere \(\times\) Group interaction. Subsequent tests showed a significant Condition \(\times\) Hemisphere interaction only in the two nonnative groups [J-High, \(F(1,22) = 16.43, p = .001\]; J-Low, \(F(1,22) = 5.665, p = .026\)]. The interaction in J-High was due to the condition effect being significant only in the right hemisphere \([F(1,22) = 9.974, p = .005]\), reflecting the right-lateralized effect of the late N400 response. In J-Low, a significant condition effect was found in both hemispheres, with a stronger effect in the right \([\text{left, } F(1,22) = 8.782, p = .007]\; \text{right, } F(1,22) = 28.18, p < .001]\).

### ERPs in the Syntactic Condition

#### ERP Morphology

Figure 4 shows the grand-average ERP waveforms in the syntactic condition (correct vs. incorrect agreement). The ERPs in the incorrect condition seem to be more negative than those in the correct condition in the left hemisphere of ENG and that of J-High in the 350–550 msec time window. In a later time window, the ERPs for the incorrect condition show a large LPC with a parietal maximum in ENG.

#### P2 Component

There was a significant group effect on the P2 amplitude at the midline sites \([F(2,61) = 7.347, p = .001]\). Planned comparisons suggested that the P2 amplitudes of J-High (6.363 ± 1.836 µV) and J-Low (6.896 ± 1.984 µV) were larger than that of ENG (4.690 ± 1.626 µV; \(p = .006\) and \(p < .001\), respectively). The Condition \(\times\) Group interaction \([F(2,61) = 3.632, p = .032]\) and the Condition \(\times\) Hemisphere \(\times\) Group interaction \([F(2,61) = 5.280, p = .008]\) were significant for the P2 amplitude at the lateral sites. Subsequent tests found a significant condition effect in ENG \([F(1,17) = .008].\)

### Table 2

Results of Omnibus ANOVAs in the Semantic Condition

<table>
<thead>
<tr>
<th>Factor</th>
<th>df</th>
<th>F</th>
<th>(\epsilon)</th>
<th>(p)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Peak latency of N400 to incongruous words</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Midline Group</td>
<td>2,61</td>
<td>3.808</td>
<td>.028</td>
<td></td>
</tr>
<tr>
<td>Midline E (\times) G</td>
<td>4,122</td>
<td>4.052</td>
<td>.004</td>
<td></td>
</tr>
<tr>
<td>Lateral Group</td>
<td>2,61</td>
<td>5.314</td>
<td>.007</td>
<td></td>
</tr>
<tr>
<td>Lateral E (\times) G</td>
<td>4,122</td>
<td>4.862</td>
<td>.002</td>
<td></td>
</tr>
<tr>
<td><strong>Peak latency of subtraction N400</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Midline Group</td>
<td>2,61</td>
<td>17.87</td>
<td>.001</td>
<td></td>
</tr>
<tr>
<td>Midline E (\times) G</td>
<td>2,61</td>
<td>18.26</td>
<td>.001</td>
<td></td>
</tr>
<tr>
<td>Lateral Group</td>
<td>2,61</td>
<td>3.497</td>
<td>.037</td>
<td></td>
</tr>
<tr>
<td>Lateral E (\times) G</td>
<td>2,61</td>
<td>4.976</td>
<td>.002</td>
<td></td>
</tr>
</tbody>
</table>

### 350–550 msec mean amplitudes

<table>
<thead>
<tr>
<th>Factor</th>
<th>df</th>
<th>F</th>
<th>(\epsilon)</th>
<th>(p)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Midline Condition</td>
<td>1,61</td>
<td>88.71</td>
<td>&lt;.001</td>
<td></td>
</tr>
<tr>
<td>Midline C (\times) E</td>
<td>2,122</td>
<td>49.91</td>
<td>.733</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Lateral Condition</td>
<td>1,61</td>
<td>106.4</td>
<td>&lt;.001</td>
<td></td>
</tr>
<tr>
<td>Lateral C (\times) E</td>
<td>5,305</td>
<td>46.86</td>
<td>.421</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Lateral H (\times) E (\times) G</td>
<td>10,305</td>
<td>4.394</td>
<td>.466</td>
<td>.001</td>
</tr>
</tbody>
</table>

### 550–850 msec mean amplitudes

<table>
<thead>
<tr>
<th>Factor</th>
<th>df</th>
<th>F</th>
<th>(\epsilon)</th>
<th>(p)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Midline C (\times) E</td>
<td>2,122</td>
<td>4.664</td>
<td>.855</td>
<td>.016</td>
</tr>
<tr>
<td>Midline C (\times) G</td>
<td>2,61</td>
<td>6.962</td>
<td>.002</td>
<td></td>
</tr>
<tr>
<td>Midline C (\times) E (\times) G</td>
<td>4,122</td>
<td>8.470</td>
<td>.855</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Lateral C (\times) H</td>
<td>1,61</td>
<td>7.613</td>
<td>.008</td>
<td></td>
</tr>
<tr>
<td>Lateral C (\times) G</td>
<td>2,61</td>
<td>5.621</td>
<td>.006</td>
<td></td>
</tr>
<tr>
<td>Lateral C (\times) H (\times) G</td>
<td>2,61</td>
<td>8.982</td>
<td>&lt;.001</td>
<td></td>
</tr>
<tr>
<td>Lateral C (\times) E (\times) G</td>
<td>10,305</td>
<td>8.069</td>
<td>.421</td>
<td>&lt;.001</td>
</tr>
</tbody>
</table>

C = Condition; H = Hemisphere; E = Electrode; G = Group; \(\epsilon\) = epsilon.
5.056, \( p = .038 \), and a significant Condition \( \times \) Hemisphere interaction in J-High \( F(1,22) = 9.653, p = .005 \), which was due to the condition effect being significant only in the left hemisphere \( F(1,22) = 4.933, p = .037 \). The condition effect found in ENG and in the left hemisphere of J-High was due to the more negative potentials in the incorrect condition, which probably index the onset of the negativity that becomes clearly visible in the 350–550 msec window.

**Mean Amplitudes: 350–550 msec**

The significant results of all omnibus ANOVAs to be presented below for the syntactic condition are listed in Table 3. No significant results involving condition were obtained for the midline sites. Instead, the ANOVA revealed a significant main effect of group. Planned comparisons showed that the mean voltages of J-High (1.944 \( \pm \) 1.785 \( \mu \)V) and J-Low (2.186 \( \pm \) 1.933 \( \mu \)V) were more positive than those of ENG (0.378 \( \pm \) 1.717 \( \mu \)V; \( p = .010 \) and \( p = .005 \), respectively); these differences are reminiscent of (and are likely to be a residue of) the group differences found in the preceding P2 component. As for the lateral sites, the main effect of condition, the Condition \( \times \) Hemisphere interaction, and the Condition \( \times \) Hemisphere \( \times \) Group interaction were significant. Subsequent ANOVAs found a significant condition effect in J-High \( F(1,22) = 5.212, p = .032 \) and a significant Condition \( \times \) Hemisphere interaction in ENG \( F(1,17) = 15.84, p = .001 \) and J-High \( F(1,22) = 18.24, p < .001 \). The interaction reflected the condition effect being significant only in the left hemisphere in ENG \( F(1,17) = 10.56, p = .005 \) and J-High \( F(1,22) = 11.05, p = .003 \), due to the more negative voltages for the incorrect condition (Incorrect \( \times \) Correct = 0.636 \( \pm \) 0.807 \( \mu \)V in ENG and \( -0.807 \pm 1.139 \mu \)V in J-High). This negativity, which starts at around the peak of the P2 component and lasts for about 300 msec, has a left fronto-central distribution in both ENG and J-High, as shown by the \( t \) maps in Figure 5.

**Mean Amplitudes: 550–850 msec**

At the midline sites, the condition main effect, the Condition \( \times \) Electrode interaction, and the Condition \( \times \) Group interaction were significant. Subsequent tests found a significant effect of condition only in ENG \( F(1,17) = 19.19, p < .001 \), reflecting the LPC for the incorrect condition (Incorrect \( \times \) Correct = 2.067 \( \pm \) 1.946 \( \mu \)V). For the lateral sites, the condition main effect,
the Condition × Hemisphere interaction, the Condition × Electrode interaction, the Condition × Group interaction, and the Condition × Electrode × Group interaction were significant. Subsequent tests found significant effects involving condition only in ENG. The condition effect in ENG \(F(1,17) = 19.88, p < .001\) reflects a positivity evoked by agreement errors (Incorrect − Correct = 1.153 ± 1.066 μV). This positivity

Figure 3. The subtraction waveforms (Incongruous − Congruous) in the semantic condition, and the peak latency and the peak amplitude of the subtraction N400 (bar diagrams). The subtraction N400 peaked earlier in ENG than in J-High and J-Low, at the midline sites and over both hemispheres. It peaked earlier in J-High than in J-Low at the midline sites and over the left hemisphere. The three groups did not differ in the peak amplitude of the subtraction N400.
Figure 4. Grand-average ERP waveforms for syntactically correct (thin lines: “Turtles move slowly.”) and incorrect words (thick lines: “Turtles moves slowly.”) at all scalp sites in three groups.
was larger over the right than the left hemisphere (Condition × Hemisphere, F(1,17) = 4.587, p = .047) and had a parietal maximum (Condition × Electrode, F(5,85) = 4.607, p = .014; see Figure 5).

DISCUSSION

ERPs were recorded from native English speakers (ENG) and two groups of native Japanese speakers with either high (J-High) or intermediate English proficiency (J-Low), while they were reading English sentences. The semantic violations elicited a large N400 response in all three groups. Group differences appeared in the time course of the brain activation in the semantic condition. The syntactic violations evoked a left-lateralized negativity in ENG and J-High (but not in J-Low) in the 350–550 msec window. A P600 component was found only in ENG. The implications of our findings for the development of second language processing at the cortical level are discussed below.

Similarities between the Groups

The general morphology of the ERP waveforms is strikingly similar across the three groups; all of them showed a large P2 (at most sites except the posterior sites), an N200 (at T5/6), and an N400 (for all types of words). This implies that the cortical processing of visually presented words embedded in a sentential context does not differ greatly between native and nonnative speakers or between nonnative speakers with different proficiency. At least at this superficial level, a view that L1 learning and L2 learning are fundamentally different processes (Bley-Vroman, 1989) is not supported. We also found that words incongruous with the preceding semantic context evoked a clear N400 component in all three groups; there were no amplitude differences between the groups either. These findings indicate that semantic context has equally robust effects across native and nonnative speakers and across nonnative speakers with different levels of proficiency.

Effects of Second Language Proficiency

We found several differences in ERPs between J-High and J-Low, who differed in English proficiency but not in the age of first exposure to English. By viewing these two groups as L2 learners at different (i.e. intermediate and advanced) developmental stages, we can infer developmental patterns in postchildhood L2 acquisition. The observed patterns will be compared with developmental patterns in child L1 acquisition below. First, the subtraction N400 in J-High peaked earlier than that in J-Low, and the effect of semantic incongruity also lasted longer in J-Low than in J-High. Because these two groups did not differ in the amplitude of the N400, an emerging generalization is that the two nonnative groups differed in the temporal profile of their brain activation but not in the strength thereof in the semantic condition. These data suggest that cortical processing of semantic information is robust from early on in the development of L2 competence, and that it only becomes faster as proficiency increases (at intermediate to late stages of development).

We also found in the semantic condition that ENG and J-High, but not J-Low, showed a hemispheric asymmetry at the anterior sites (regardless of semantic congruity). The absence of the asymmetry in J-Low was attributable to a pronounced right anterior negativity (cf. Hahne & Friederici, 2001). L2 processing has sometimes been associated with more right hemisphere involvement, and one may take the right anterior negativity in J-Low as additional support for such a claim. However, a hemispheric asymmetry at the anterior sites appeared in J-High, suggesting that the presence of the right anterior negativity is not a categorical marker of L2 processing, but rather reflects the low proficiency of L2 speakers. Thus, our data indicate that lateralization to the left hemisphere gets stronger as L2 proficiency increases (Perani et al., 1998).

In the syntactic condition, S–V number agreement errors elicited a left-lateralized negativity in the 350–550 msec latency range in ENG and J-High but not in J-Low, whose scalp distribution (Figure 5) suggests

| Table 3. Results of Omnibus ANOVAs in the Syntactic Condition |
|------------------|------------------|------------------|------------------|
| Factor           | df              | F               | p               |
| 350–550 msec mean amplitudes |  |  |  |
| Midline Group    | 2,61            | 5.387           | .007            |
| Lateral Condition| 1,61            | 4.690           | .034            |
| C × H            | 1,61            | 30.61           | <.001           |
| C × H × G        | 2,61            | 3.233           | .046            |
| 550–850 msec mean amplitudes |  |  |  |
| Midline Condition| 1,61            | 4.771           | .003            |
| C × E            | 2,122           | 4.742           | .015            |
| C × G            | 2,61            | 8.533           | .001            |
| Lateral Condition| 1,61            | 5.420           | .023            |
| C × H            | 1,61            | 9.952           | .002            |
| C × E            | 5,305           | 3.095           | .037            |
| C × G            | 2,61            | 6.834           | .002            |
| C × E × G        | 10,305          | 2.414           | .039            |

C = Condition; H = Hemisphere; E = Electrode; G = Group; ε = epsilon.
that it has some affinity with the LAN (Münnte et al., 1993). The timing of this negativity fits with the second stage of the three-stage model of language comprehension proposed by Friederici (1995), at which thematic and semantic processing are assumed to take place. Hence, we interpret the negativity found in ENG and J-High as reflecting a failure in checking agreement information (on the subject and the verb) which can guide thematic role assignment. In contrast to the N400, this negativity was missing in J-Low. Because J-High showed the negativity whose scalp distribution is similar to the one found in ENG, we argue that given high enough proficiency, cortical processing of agreement information can become native-like even in late L2 learners, although its development is determined by proficiency more profoundly than is the development of semantic processing. One counter-argument against this conclusion would be that the difference between the semantic and the syntactic condition was confounded by the word category of critical words (nouns vs. verbs). Future research controlling this factor is desirable.

Overall, these ERP findings suggest that processing of an L2 acquired after childhood approximates that of an L1 as proficiency increases, which is in disagreement with the critical period hypothesis claiming that postchildhood L2 learning is a fundamentally different process from child L1 learning. Moreover, the developmental pattern inferred from the differences between J-High and J-Low, namely, the early development of semantic processing compared with syntactic processing, seems to be similar to developmental patterns in L1 learning reported elsewhere. For example, Friederici and Hahne (2001) suggested that in child L1 learning, the N400 becomes adult-like earliest, followed by the ELAN and the P600, and Hahne, Eckstein, and Friederici

**Figure 5.** These t maps show the statistical effects of subject–verb (S-V) agreement errors in the syntactic condition. ANOVAs indicate that the verbs disagreeing in number with the subject elicited more negative potentials from around the peak of the P2 component (i.e., around 250 msec) until around 550 msec after stimulus onset in ENG and J-High. From 550 msec onward, the incorrect verbs evoked a late positive component (P600) only in ENG. To further explore the nature of these effects, t values were obtained from paired t tests comparing the correct and the incorrect condition at each scalp site in each group in two time windows: 250–550 msec and 550–850 msec. Contour maps of these values were created using the spline interpolation. The white lines in the maps show thresholds for statistical significance; from the outmost line, they correspond to p = .05, .025, .01, .0075, .005, .0025, .001, respectively (uncorrected for multiple comparisons). The maps show that the early negativity elicited by S-V agreement errors in ENG and J-High has a left fronto-central distribution, whereas the P600 found in ENG has a centro-parietal distribution with right hemisphere dominance.
differences indicate that nonnative speakers may not argue for the critical period hypothesis, because such nonnative speakers as a whole (J-High and J-Low) would qualitative differences in ERPs between ENG and the Our nonnative subjects included high-proficiency speak- like brain responses). Future research is awaited here a long stay led to high proficiency (and to more native- whether high proficiency made a long stay possible or the length of stay abroad as well, and it is not clear that early starters are likely to become more proficient in the L2 than late starters. Future research should test the possibility that the electrophysiological differences found between native and nonnative speakers (Hahne, 2001; Hahne & Friederici, 2001; Weber-Fox & Neville, 1996). Similar results may have similar causes. It is usually the case that nonnative speakers are less proficient in the target language than native speakers of that language, and that early starters are likely to become more proficient in the L2 than late starters. Future research should test the possibility that the electrophysiological differences found between native and nonnative speakers (Hahne, 2001; Hahne & Friederici, 2001; Weber-Fox & Neville, 1996, 2001) are attributable to differences in proficiency. A more diffi- cult question is what causes a proficiency difference in the first place. For example, J-High and J-Low differed in the length of stay abroad as well, and it is not clear whether high proficiency made a long stay possible or a long stay led to high proficiency (and to more native-like brain responses). Future research is awaited here as well.

Native/Nonnative Contrasts

Our nonnative subjects included high-proficiency speak- ers (J-High) as well as low-proficiency ones (J-Low), and qualitative differences in ERPs between ENG and the nonnative speakers as a whole (J-High and J-Low) would argue for the critical period hypothesis, because such differences indicate that nonnative speakers may not become native-like in terms of brain responses even if they achieve high-proficiency. Hence, it is discussed below whether the differences found between ENG and the nonnative groups are of qualitative or quanti- tative nature. First, we found in the syntactic condition that the peak amplitude of the P2 component at the midline sites was larger in J-High and J-Low than in ENG (cf. Weber-Fox & Neville, 2001). In L1 processing, the P2 component evoked by low-frequency words is larger than that evoked by high-frequency words (King & Kutas, 1998). We speculate that the P2 of nonnative speakers was larger than that of native speakers be- cause the former group had less frequent contact with English than the latter. Crucial here is that frequency is a quantitative factor; we argue that the difference in the P2 amplitude does not index a qualitative difference between native and nonnative speakers. Another quantitative difference between native and nonnative speakers was in the peak latency of the N400 response to incongruous words; it was earlier for the native controls than for the two nonnative groups. The dura- tion of the N400 was also longer in the nonnative groups than in the native controls; such a difference has been found between children and adults in L1 processing (Friederici & Hahne, 2001). These differences are clearly of quantitative nature, and do not constitute evidence for qualitative discontinuity between native and nonnative processing.

Candidates for qualitative differences between native and nonnative speakers were also found. The peak of the N400 response to incongruous words became later towards the parietal sites in the nonnative groups but not in ENG. An inspection of the individual data in the nonnative groups suggested that the frontal peak and the parietal peak were of two overlapping but differ- ent components. Pytlkänen and Marantz (2003) claim that the electrophysiological N400 can be decomposed into maximally three subcomponents, which can be distinguished by magnetoencephalography. Also, the early part of the N400 evoked by semantic incongruity is frontally distributed, whereas the later part has a centro-parietal maximum (Niedeggen, Rössler, & Jost, 1999), and the presence of “anatomically distinct N400s” in the same time window is suspected in the ERP literature (Kutas & Federmeier, 2000). Taking into account the decomposition of the N400, we interpret the Electrode × Group interaction observed in the N400 peak latency as follows: only the later, more posterior subpart of the N400 reflecting semantic integration processes (Niedeggen et al., 1999) is delayed in J-High and J-Low compared to ENG, and the three groups did not differ in the early, anterior subcomponent indexing lexical activation (Pytlkänen & Marantz, 2003; Niedeggen et al., 1999). In other words, what seemed to be a qualitative difference between the native and nonnative groups may rather be a quantitative difference (i.e., latency delay), in the light of a fine-grained account of the N400 proposed elsewhere for independent reas-

(2004) recently reported that the ELAN is absent in 6-year-olds, whereas a late N400 is present. ERP studies of native language processing by adults also indicate that semantic processing as evidenced by the N400 is more robust than syntactic processing, electrophysio- logical correlates of which (e.g., LAN, P600) are some- times absent even in healthy subjects (Osterhout, 1997). Hence, the early maturation of semantic processing in L2 compared with syntactic processing found in our study may simply mirror the normal properties of the brain which also constrain L1 learning and processing. This conclusion is in line with a theoretical argument provided by Epstein et al. (1996), and is supported by behavioral studies (Vainikka & Young-Scholten, 1996; Radford, 1990) which report similar developmental patterns across L1 and L2, in which “content words,” which mainly carry semantic information, appear early in the learner’s speech and “function words” and gram- matical morphemes, which mainly carry syntactic infor- mation, appear late.

Our data above provide yet another implication for bilingual ERP research. In this study, the difference in L2 proficiency appeared as a latency difference in semantic processing but in syntactic processing, it appeared as the presence/absence of a component. Previous ERP studies have shown that the age of first exposure and the native/nonnative contrast also have such different effects on semantic and syntactic processing (Hahne & Friederici, 2001; Weber-Fox & Neville, 1996). Similar results may have similar causes. It is usually the case that nonnative speakers are less proficient in the target language than native speakers of that language, and that early starters are likely to become more proficient in the L2 than late starters. Future research should test the possibility that the electrophysiological differences found between native and nonnative speakers (Hahne, 2001; Hahne & Friederici, 2001) and between early and late bilinguals (Weber-Fox & Neville, 1996, 2001) are attributable to differences in proficiency. A more diffi- cult question is what causes a proficiency difference in the first place. For example, J-High and J-Low differed in the length of stay abroad as well, and it is not clear whether high proficiency made a long stay possible or a long stay led to high proficiency (and to more native-like brain responses). Future research is awaited here as well.
ons. In the 550–850 msec time window, we found an LPC elicited by incongruous words only in ENG. This component was statistically weak even in ENG, and is not present in children aged 5–11 (Juottonen, Revonsuo, & Lang, 1996), suggesting that it is not so stable even in native language processing. Thus, the absence of an LPC in the L2 groups may simply reflect the normal course of language development.

S–V number agreement errors elicited a P600 with a centro-parietal maximum only in ENG (cf. Osterhout & Mobley, 1995). In contrast to the LPC elicited by semantic violations (Juottonen et al., 1996), a P600 was found in children (Friederici & Hahne, 2001). Thus, the absence of the P600 to S–V agreement errors in the nonnative groups may reflect a true qualitative difference between native and nonnative speakers. However, other studies found a P600 elicited by syntactic errors in nonnative speakers (Sabourin, 2003; Hahne, 2001). A possible account of the lack of the P600 in our study is that agreement errors were not salient enough to evoke a P600 in L2 speakers, but this account is not supported by our behavioral data: J-High performed slightly better at detecting agreement errors than ENG. Another possibility is suggested by Sabourin’s (2003) study, which reported that mismatches in gender features in L2 Dutch produced a clear P600 in speakers of German, which has a different gender system to Dutch, but not in speakers of languages in which gender is differently expressed (Romance languages such as French) or not expressed (English). Theoretical linguists argue that Japanese has a system of agreement itself but not features such as person and number which induce syntactic S–V agreement (Fukui & Sakai, 2003; Harada, 1976). Hence, we interpret the absence of the P600 in our Japanese subjects as resulting from the absence of agreement-inducing features in Japanese. The lack of the P600 suggests that the cognitive processes reflected by the P600 (e.g., on-line syntactic reanalysis/repair as proposed by Hahne & Friederici, 1999) cannot be triggered by syntactic features acquired after a critical period. This circumscribed impairment in L2 processing, we propose, is a true qualitative difference from native language processing (cf. Friederici, Steinhauer, & Pfeifer, 2002). One caveat is that it remains an open question whether a late P600 appeared in J-High well after the analysis period we used (i.e., 900 msec after stimulus onset).

Semantic incongruity had prolonged effects in the right hemisphere of J-Low and J-High, and one may take this as evidence for more right hemisphere involvement in L2 processing. However, the generator of the N400 is likely to be in the left hemisphere (Kutas, Hillyard, & Gazzaniga, 1988), and the N400 in L1 processing itself often has a right hemisphere dominance (Kutas, Van Petten, & Besson, 1988); the peak-to-peak amplitude of the N400 to incongruous words in our study was also larger in the right hemisphere regardless of group. In the semantic condition, analyses of the P2 amplitude at the lateral electrodes found that only in J-High did the Hemisphere × Electrode interaction reach significance, for which we have no explanation.

Conclusion

Our ERP data indicate that the cortical processing of a second language approximates that of a first language as proficiency increases. The process of such approximation was argued to be similar to the development of first language competence in the brain. Hence, in line with recent behavioral and hemodynamic studies, we obtained little evidence for an absolute critical period effect and discontinuity between L1 and L2 learning. We single out the lack of a P600 triggered by syntactic features acquired after childhood as a true point of divergence between native speakers and postchildhood L2 learners. In other respects, a strong position that argues for an L1–L2 identity seems to be tenable.

METHODS

Subjects

Two groups of late learners of L2 English whose first language was Japanese (J-High and J-Low) and a group of native speakers of English (ENG) took part in the experiment, after providing written content. All participants had normal or corrected-to-normal vision, and had no record of psychological or neurological disorders. All the Japanese participants were right-handed but the native control group included three left-handers. The 18 native English speakers (8 men; mean age 27.7 ± 7.5) were from either Australia, Canada, New Zealand, UK, or USA, and were temporarily staying in Japan. For all the Japanese participants, English was their second language. All but three started learning English at age 12. The remaining three started at age 10 or 11. The group with high English proficiency (J-High) consisted of 23 advanced L2 speakers of English who daily used English professionally in Japan (11 men; mean age 29.8 ± 3.25). All of them had lived in an English-speaking country after adolescence (mean length of stay 3.75 years) (Figure 6), and had very high scores on the TOEIC Test (mean 925.1/990 ± 23.9, range 870–985). The other group of Japanese subjects (J-Low) consisted of 23 intermediate L2 learners of English (11 men; mean age 28.3 ± 3.71) whose English proficiency was significantly lower than that of J-High (mean TOEIC score 632.3/990 ± 87.8, range 395–767). Two of them used English daily in their professions. J-High received formal instruction of English for significantly longer than J-Low [t(44) = 2.569, p = .014], and J-High rated their own proficiency in English more highly than did J-Low [F(1,44) = 79.92, p < .001] (Figure 6).
Both groups rated their proficiency in comprehension (listening and reading) more highly than that in production (speaking and writing) \( [F(1,44) = 36.92, p < .001] \).

Stimuli

Because a preliminary behavioral experiment revealed that L2 word recognition is severely affected by quantitative word properties (length and frequency), we decided that only three or four letters long, frequent words be used as “critical words” in the ERP experiment. Sixty nouns (187.4 ± 119.5 counts in the corpus of Kučera & Francis, 1967) and 60 verbs (212.4 ± 154.8 counts) were selected as critical words for the semantic and syntactic conditions, respectively. A total of 120 pairs of stimulus sentences consisting of basic English words were devised (60 pairs for the semantic and another 60 for the syntactic condition) (see Appendix). Paired sentences shared the same sentence frame. In the semantic condition, one sentence in each pair was semantically anomalous by virtue of one word that was taken from another sentence frame (Congruous: “The house has ten \textit{rooms} in total.” vs. Incongruous: “The house has ten \textit{cities} in total.”). The same critical word was used as an incongruous word in one pair once and as a congruous word in another pair once. In the syntactic condition, one sentence in each pair contained an S–V agreement error, whereas the other did not (Correct: “Turtles \textit{move} slowly.” vs. Incorrect: “Turtles \textit{moves} slowly.”). Agreement errors were used for the syntactic condition because they can be embedded in simple active sentences not inducing much integration difficulty on nonnative speakers (cf. Hahne & Friederici, 2001) and are purely syntactic in nature unlike phrase structure violations (Osterhout & Mobley, 1995). The critical words always appeared in sentence-internal positions rather than in sentence-final positions (Osterhout, 1997). Repetition of the same content words was avoided, except for several phrases (e.g., \textit{every day}) that appeared at the sentence-final position. The sentences were 5 to 11 words long (average 7.2 ± 1.1) in the semantic and 3 to 9 words long (average 5.3 ± 1.34) in the syntactic condition. Each of these sentences was divided into three to five chunks of one to three words (Appendix). Each critical word constituted a single chunk. Two lists of 120 sentences were created; each list contained one sentence from each of the 120 sentence pairs. A half of the subjects were shown one list and the other half the other list. Thus, each subject read 120 sentences consisting of 30 congruous, 30 incongruous, 30 S–V agreeing, and 30 S–V disagreeing sentences. These sentences were presented in a pseudorandom order.

Procedure

The subject was seated on a comfortable chair in a shielded room. Following a fixation mark that lasted for 2 sec, the stimulus sentence appeared on a chunk-by-chunk basis (not on a word-by-word basis) in the center of a computer monitor. To reduce onset responses from the primary visual cortex, the 600-msec interstimulus intervals were filled by random-dot patterns (refreshing rate 30 Hz), which appeared in the rectangular space where the subsequent chunk was about to be shown. The duration of a chunk that consisted of a
critical word was set at 850 msec. The durations of other chunks varied as a function of their lengths [e.g., London (850 msec) is the (925) biggest (875) city (850) in Britain. (1125)]. The subjects were instructed to read the sentences for comprehension silently, without blinking. When one sentence was over, a prompt for a button press appeared on the screen, which remained until the subject pressed the button. The subject used the button just to proceed to the next sentence. The subject’s comprehension was checked by asking easy T/F questions about the content of the sentences. Each T/F question was printed on a piece of paper and was shown to the subject immediately after the appropriate sentence. The subjects responded verbally. One fourth of all stimulus sentences were followed by a T/F question. Before starting, the subjects were given a practice session. After the ERP session finished, the subjects were given an off-line acceptability judgment task. They were shown exactly the same sentences as those they saw in the ERP session and were asked to mark the sentences that they thought were not appropriate for any reasons. Off-line judgments could not be obtained from three English-speaking subjects due to time constraints.

**Data Analysis**

The subjects’ responses to the T/F questions and in the off-line acceptability judgment task were analyzed in terms of accuracy. The data from the T/F questions were analyzed by a one-way ANOVA with Group (ENG vs. J-High vs. J-Low) as the factor. The semantic and the syntactic part of the acceptability judgments were analyzed separately, in a Condition (congruous vs. incongruous meaning, or correct vs. incorrect agreement) × Group design. The EEGs were baseline-adjusted relative to the mean voltage in the prestimulus 100-msec interval. The ERP responses to all critical words were analyzed whether or not they were accurately judged in the off-line judgment task. In both the semantic and syntactic conditions, the peak latency and peak amplitude of the P2 (at all sites except P3/4 and T5/6) and the mean amplitudes in the 350–550 msec and 550–850 msec time windows at all sites were analyzed by ANOVAs. The semantic and the syntactic conditions were analyzed separately. The data from the midline sites were analyzed in a Condition × Electrode × Group design, whereas those from the lateral sites were analyzed in a Condition × Hemisphere (left vs. right) × Electrode × Group design. In the semantic condition, the peak latency and peak-to-peak amplitude of the N400 response to incongruous words, and the peak latency and peak amplitude of the subtraction N400 at the three midline sites (Fz, Cz, Pz) and six lateral sites (F3/4, C3/4, P3/4) were measured additionally. The measurements at the midline sites were analyzed in an Electrode × Group design, and those at the lateral sites were analyzed in a Hemisphere × Electrode × Group design. All factors except Group were analyzed as repeated-measures factors. Significant interactions involving condition in these omnibus ANOVAs led to further tests of condition at lower levels (e.g., at each electrode or in each group). Significant main effects of group in ANOVAs led to planned comparisons among the three groups. Huynh and Feldt’s (1976) correction was used if the sphericity assumption was violated. Normalization to assess topographical differences was not applied (Urbach & Kutas, 2002). The significance level was set at .05; only the significant or marginally significant results were reported in the Results section. The peak latency and peak amplitude of the N200 component at T5/6 were measured but yielded no significant results involving group.

**APPENDIX**

Each stimulus sentence in the ERP experiment was divided into chunks of words, and the participant was shown one chunk after another on the computer monitor. The list of the stimulus sentences presented below contains slashes (/) at the boundaries of the chunks. The critical words that triggered EEG acquisition are italicized. The words in the round brackets (…) in the semantic condition (A1 below) are semantically incongruous with the preceding context. Those in the syntac-
tic condition (A2 below) are verbs that do not agree in number with the subject.

### A1. Semantic Condition

1. She slept / in my / bed (*law) / that night.
3. She has / much fat / in the upper / arms (*art) / and the hips.
4. The gallery / exhibits / pop / art (*arms) / now.
5. Scuba diving / in the / clear blue / sea (*job) / is great.
7. There are / no clouds / in the / sky (*box) / today.
8. Leave / your key / in this / box (*sky) / please.
9. My neighbors / feed / this / dog (*cup) / every day.
10. Pour / some tea / in this / cup (*dog) / please.
12. I will / measure / your / size (*test) / accurately.
13. We have / three / unmarried / sons (*lines) / now.
15. The ice-skater / broke / his / leg (*road) / last year.
16. A truck is / blocking / the / road (*leg) / now.
17. Please / spell / your / name (*face) / clearly.
18. Wash / your / face (*name) / in the / morning.
19. He edited / only / one / book (*wife) / this year.
22. The factory / is in the / industrial / area (*girl) / there.
23. The restaurant / serves / low-fat / food (*fire) / now.
25. This cheese / is made / from / milk (*ships) / and cream.
26. The typhoon / sank / the / ship (*milk) / right away.
27. The actress / appeared / in this / film (*foot) / last year.
28. The athlete / damaged / his / right / foot (*film) / last week.
29. She finished / first / in the / race (*wall) / today.
30. Hang / this mirror / on the / wall (*race) / please.
31. I noticed / tears / in his / eyes (*air) / afterwards.
32. The moon / doesn’t / have / air (*eyes) / at all.
33. Rabbits / have long / ears (*lakes) / and a / short tail.
34. Fishing / in this / lake (*ear) / is fun.
35. The helicopter / landed on / top (*age) / of the / building.
36. The profile / omits / the actor’s / age (*top) / suspiciously.
37. The earth / orbits the / sun (*gun) / once / a year.
38. It’s illegal / to possess / this / gun (*sun) / here.
39. The executive / drives / an expensive / car (*park) / these days.
40. There’s / a zoo / in the / park (*car) / there.
41. There are / exams / at the / end (*war) / of the term.
42. America / fights a (an) / war (*end) / against / terrorism.
43. The military / uniform / includes a / bat (*boy) / as well.
44. The boss / hired a / boy (*hat) / as an / assistant.
45. The motorcycle / slipped / on / oil (*news) / here.
46. The media / report / bad / news (*oil) / every day.
47. The story is / based on / fact (*home) / to some / degree.
48. The guys / left / my / home (*fact) / yesterday.
49. Remember / to lock / that / door (*hand) / please.
50. The captain / shook / hands (*doors) / with the / referee.
51. My colleague / stole / my / idea (*head) / once.
52. The passenger / injured / his / head (*idea) / severely.
53. The house / has / ten / rooms (*cities) / in total.
54. London / is the / biggest / city (*room) / in Britain.
55. Robberies / occur / in this / town (*body) / everyday.
56. The therapist / massages / my / body (*town) / very well.
57. That Italian / drank / the red / wine (*king) / slowly.
58. The royal / family / obeyed the / king (*wine) / faithfully.
60. The victims / suffered / a financial / loss (*song) / as well.

### A2. Syntactic Condition

2. Candidates / wait (*waits) / for the / results / anxiously.
3. Earthquakes / hit (*hits) / this / island / every year.
4. New / members / join (*joins) / the group / every year.
5. Sales / keep (*keeps) / increasing / these days.
6. These / chairs / look (*looks) / nice.
7. Customers / read (*reads) / the instructions / only / briefly.
8. The employers / know (*knows) / Ken’s / talent.
9. Two pianists / pick (*picks) / up / this award / every year.
10. These / shops / sell (*sells) / gold.
11. These / symbols / mean (*means) / nothing / to me.
12. Weak / teams / win (*wins) / occasionally.
13. These / cakes / cost (*costs) / 500 yen / each.
14. All my / friends / love (*loves) / soccer.
15. Some / scientists / find (*finds) / solutions / by chance.
16. Plants / grow (*grows) / from / seeds.
17. Pretty / flowers / fill (*fills) / my yard / in spring.
18. These / travelers / want (*wants) / a map.
20. Many / students / hope (*hopes) / to study / abroad.
22. Fans / talk (*talks) / about / their / heroes.
24. My cats / play (*plays) / with / a ball.
27. My parents / buy (*buys) / only / cheap / stuff.
28. Guides / show (*shows) / us / where / to visit.
29. Forests / need (*needs) / water.
30. Many / people / die (*dies) / in / traffic / accidents.
32. Important / guests / stay (*stays) / at / this / hotel.
33. Politicians / meet (*meets) / here / unofficially.
34. These / trains / run (*runs) / very fast.
35. These / banks / plan (*plans) / to / merge / this year.
36. Patients / call (*calls) / me / Dr. Bush.
37. Tourists / take (*takes) / pictures / everywhere.
38. Some / poisons / kill (*kills) / even / elephants.
39. These / knives / cut (*cuts) / well.
40. Positive / attitudes / lead (*leads) / to / success.
41. Buses / stop (*stops) / here.
42. These / companies / pay (*pays) / well.
43. Some trees / live (*lives) / for more than / a hundred / years.
44. Many / couples / come (*comes) / to / this / museum.
45. My sisters / work (*works) / at / the / same / store.
46. Some / adults / lack (*lacks) / vitamin D.
47. Some / ladies / walk (*walks) / to / lose / weight.
49. Researchers / bold (*holds) / a / conference / here / every year.
50. Turtles / move (*moves) / slowly.
51. Dreams / turn (*turns) / into / reality.
52. Discount / tickets / save (*saves) / you / some money.
53. My kids / ask (*asks) / me for / a / dollar / every day.
54. Nurses / tell (*tells) / me / to / be / quiet.
55. The rules / say (*says) / that / it / is / permitted.
56. Mothers / care (*cares) / about / their / babies.
57. Many / firms / send (*sends) / direct / mail.
58. These / sections / open (*opens) / at / 11.
59. Doctors / help (*helps) / ill / people.
60. Armies / use (*uses) / weapons.

Acknowledgments

We thank Osamu Nagata and Yasuyuki Takeshima for technical assistance, Aya Ihara and Yohei Tamura for helping the experiment, and Cheryl Lamers and Mitsue Ogasawara for recruiting subjects. Hiroyo Hagiwara, Roger Hawkins, Hiromu Sakai, and three anonymous reviewers provided valuable comments. All remaining errors are ours.

Reprint requests should be sent to Shiro Ojima, Department of Integrative Physiology, National Institute for Physiological Sciences, Myodaiji, Okazaki, 444-8585, Japan, or via e-mail: sojima@nips.ac.jp.

Notes
1. One may argue that the processes reflected by late positive components are, in general, impaired in L2 processing, as both the LPC following the N400 and the P600 following the early left-lateralized negativity were missing in J-High and J-Low. However, it remains unclear whether these late positive components are one and the same component or not (Frisch, Kotz, von Cramon, & Friederici, 2005; Coulson et al., 1998).
2. TOEIC stands for Test of English for International Communication. According to a survey of the scores from 1st to 98th TOEIC “Secure Program” tests published by The Institute for International Business Communication, only 2.4% of the total test takers scored 895 or higher.
3. To determine the duration of a particular chunk, we used a function that takes as variables the number of words in a chunk and the length of each of these words and yields the duration of the chunk. The coefficients in this function were determined so that the speed of chunk presentation would not be too fast for nonnative speakers of English.

REFERENCES


