# Second-language Instinct and Instruction Effects: Nature and Nurture in Second-language Acquisition

**Noriaki Yusa[1,2], Masatoshi Koizumi[2,3], Jungho Kim[3], Naoki Kimura[3],
Shinya Uchida[4], Satoru Yokoyama[3], Naoki Miura[3,5],
Ryuta Kawashima[3], and Hiroko Hagiwara[2,6]**

## Abstract

■ Adults seem to have greater difficulties than children in acquiring a second language (L2) because of the alleged "window of opportunity" around puberty. Postpuberty Japanese participants learned a new English rule with simplex sentences during one month of instruction, and then they were tested on "uninstructed complex sentences" as well as "instructed simplex sentences." The behavioral data show that they can acquire more knowledge than is instructed, suggesting the interweaving of nature (universal principles of grammar, UG) and nurture (instruction) in L2 acquisition. The comparison in the "uninstructed complex sentences" between post-instruction and pre-instruction using functional magnetic resonance imaging reveals a significant activation in Broca's area. Thus, this study provides new insight into Broca's area, where nature and nurture cooperate to produce L2 learners' rich linguistic knowledge. It also shows neural plasticity of adult L2 acquisition, arguing against a critical period hypothesis, at least in the domain of UG. ■

## INTRODUCTION

A significant feature of language acquisition is that children, in a relatively short time, acquire a language while attaining a rich knowledge of their language which surpasses their actual experiences (Chomsky, 1980). A large discrepancy between the limited input that children are exposed to and the rich linguistic knowledge that they acquire poses what is called the poverty-of-the-stimulus (PoS) argument (Chomsky, 1980, 2000): how children end up knowing more than they experience. This problem has been at the heart of a heated debate since the inception of cognitive science (Carruthers, Laurence, & Stitch, 2005; Reali & Christiansen, 2005; Samuels, 2004; Tomasello, 2003; Chomsky, 1980, 2000; Elman et al., 1996; Pinker, 1994). Finding an answer to the problem has engendered a wide variety of research programs. One of the solutions based on theories of biolinguistics (i.e., generative grammar) is to postulate and characterize an innate language faculty known as the language instinct or universal grammar (UG) in human brains, which provides children with a set of principles for developing a grammar on the basis of limited linguistic input (Chomsky, 2000; Crain & Thornton, 1998; Pinker, 1994). Thus, the PoS argument furnishes the most compelling rationale for the presence of UG at least in first-language (L1) acquisition.

In sharp contrast to the remarkable feat of L1 acquisition, the language instinct with which humans are endowed seems to deteriorate or dysfunction with age in L2 acquisition beyond a hypothesized critical period (DeKeyser, 2000; Pinker, 1994; Bley-Vroman, 1990; Johnson & Newport, 1989). Of special interest in this connection is whether the language instinct or UG is still in place and remains operative in adult L2 acquisition, functioning as the so-called second-language instinct (White, 2003; Schwartz, 1998). If L2 acquisition were fundamentally different from L1 acquisition, L2 grammars might not conform to principles of UG that constrain L1 grammars, resulting in "wild" grammars that do not exhibit properties of natural languages. If L2 learners, however, prove to have developed abstract knowledge underdetermined by their L2 input and their L1, this PoS argument strongly suggests that the innate language instinct is not "dismantled" but mediates L2 acquisition as a second instinct, thereby enabling L2 learners to know more than they experience.

The purpose of the present study was to investigate whether L2 learners' knowledge would go beyond the input or stimuli that they had received during instruction, by examining the acquisition of a syntactic rule called negative inversion (NI): Negative adverbs (*never*, *seldom*, *rarely*, etc.), when placed at the beginning of a sentence, obligatorily trigger inversion and must be followed by auxiliaries (*be*-verbs, *can*, *must*, *may*, etc.): *I will never eat sushi* → *Never will I eat sushi*. Note that in contrast

---

[1]Miyagi Gakuin Women's University, Sendai-shi, Japan, [2]Research Institute of Science and Technology for Society (RISTEX), Tokyo, Japan, [3]Tohoku University, Sendai-shi, Japan, [4]National Center of Neurology and Psychiatry, Kodaira-shi, Japan, [5]Kochi University of Technology, Kami-shi, Japan, [6]Tokyo Metropolitan University, Tokyo, Japan

to the grammatical sentence involving a positive adverb, *Often I eat sushi*, the noninverted sentence involving a negative adverb *Never I will eat sushi*, although semantically intelligible, is syntactically incorrect, which clearly shows that NI is a purely syntactic rule.

Although words in a sentence are arranged in a linear order like beads on a string (Figure 1A), sentences are assumed to have abstract hierarchical structure (Hauser, Chomsky, & Fitch, 2002; Chomsky, 1957). For example, the same linear sequence of words, *Japanese history teacher*, can either mean a Japanese person who teaches history, represented as the bracketed structure [Japanese [history teacher]] (Figure 1B), or a teacher who teaches Japanese history, represented as [[Japanese history] teacher] (Figure 1C). This simplex example clearly shows that our mental computation depends on the hierarchical structure, which is referred to as the principle of structure dependence, one of the basic principles of UG. More specifically, our mental grammar does not refer to linear or structure-independent notions such as "first word" or "leftmost word," but exclusively to hierarchical or structure-dependent notions such as "subject" or "matrix clause" (Chomsky, 1980). The principle of structure dependence puts severe limits on logically possible rules that human languages exploit: Not all conceivable rules are realized in human languages (Moro, 2008).

With respect to structure dependence in L2 acquisition, the study of Japanese adults learning NI can offer a critical test of whether UG is still accessible in L2 acquisition for the following three reasons: First, the Japanese language does not have NI, which makes it unlikely that their L2 knowledge of NI is directly transferred from Japanese grammar or acquired by natural analogical extension; second, NI is acquired late in L1 (Slobin, 2003) and is considered to be quite difficult for L2 learners to acquire, due to the relatively low frequency of occurrence of the structure in native English input and to being primarily a structure found in literary genres, thereby making it unlikely that L2 learners would be able to attain knowledge of NI strictly on the basis of general language input; third, mastery of NI requires purely syntactic knowledge, specifically knowledge of structure dependence. Simplex sentences such as *I will never eat sushi* are consistent with both a structure-dependent rule (i.e., move the main clause auxiliary after a fronted negative adverb) and a linear structure-independent rule or a counting strategy (i.e., move the first or leftmost auxiliary after a fronted negative adverb), correctly changing into *Never will I eat sushi*.

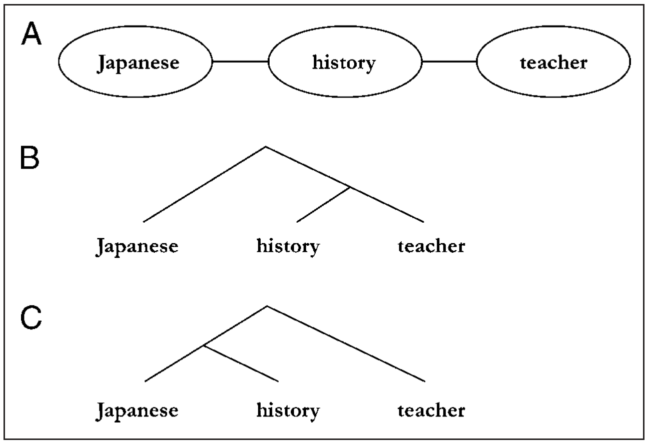

**Figure 1.** A common-sense linguistics assumes that words are strung together one after another like beads on a string or links in a chain (A), which fails to reflect native-speaker's intuitions about what words function as units. On the other hand, the hierarchical structures (B, C) allow a straightforward explanation of the ambiguity of "Japanese history teacher."

**Table 1.** Structure-dependent Rule versus Structure-independent Rule

---

(1) a. I will *never* eat sushi
b. *Never* **will** I ___ eat sushi

(2) a. [Those students who will fail a test] are *never* hardworking in class
b. * *Never* **will** [those students who _____ fail a test] are hardworking in class
c. *Never* **are** [those students who will fail a test] ____ hardworking in class

(3) a. Those restaurants are *rarely* full of [people who have been there before]
b. *Rarely* **are** those restaurants ____ full of [people who have been there before]
c. * *Rarely* **have** those restaurants are full of [people who ___ been there before]

---

The position marked by "___" indicates the original position of the fronted auxiliary in bold and "*" indicates ungrammaticality. (1b) can be formed from (1a) by either a structure-dependent rule [i.e., moving the main clause auxiliary (structurally highest auxiliary) after the fronted adverb] or a structure-independent rule (i.e., moving the first or leftmost occurrence of the auxiliary after the fronted negative adverb). The structure-dependent rule successfully produces (2c) from (2a), whereas the structure-independent rule wrongly produces (2b) from (2a). Note that (3b) can be derived from (3a) by either a structure-dependent rule or a structure-independent rule, where the first or leftmost auxiliary (*are*) happens to be the main clause auxiliary. To work in all cases, the principle of structure dependence moves the hierarchically highest auxiliary from the corresponding declarative sentences rather than the leftmost or first auxiliary in the linear order.

Note here that a general model of inductive bias learning prefers the "simpler" structure-independent rule. However, the structure-independent rule fails in complex sentences involving relative clauses. For example, with the complex sentence, *Those students who will fail a test are never hardworking in class*, the structure-independent rule would move the first auxiliary *will* and give the ungrammatical result: *Never will those students who fail a test are hardworking in class*. The structure-dependent rule, on the other hand, moves the main clause auxiliary (i.e., *are*) after the fronted negative adverb *never* and generates the grammatical sentence: *Never are those students who will fail a test hardworking in class*. To explain the ungrammaticality of (2b) in Table 1, we need to distinguish between an auxiliary embedded in the relative clause (*will*) and the main clause auxiliary (*are*). This involves assigning an abstract structure to the declarative sentence in (2a) in Table 1. To work in all cases, NI needs to move the hierarchically highest auxiliary from the corresponding declarative sentence rather than the first auxiliary in the linear order (Figure 2). The rule of NI, similar to every syntactic rule of human languages, is structure-dependent.

It may be objected that adult L2 learners simply come across the alternations such as, *Those students who will fail a test are never hardworking in class → Never are those students who will fail a test hardworking in class.* To our knowledge, however, no English-as-a-Foreign-Language (EFL) textbook refers to the notion of structure dependence, or lists complex sentences involving relative clauses in the explanation of NI, which makes it unlikely that L2 learners could attain the knowledge of complex sentences with NI from a sufficient number of examples in their textbooks. It is claimed that human languages make exclusive use of structure-dependent rules that require an analysis of a sentence into an abstract hierarchical structure and operate on abstract grammatical constructs such as phrases or clauses, instead of structure-

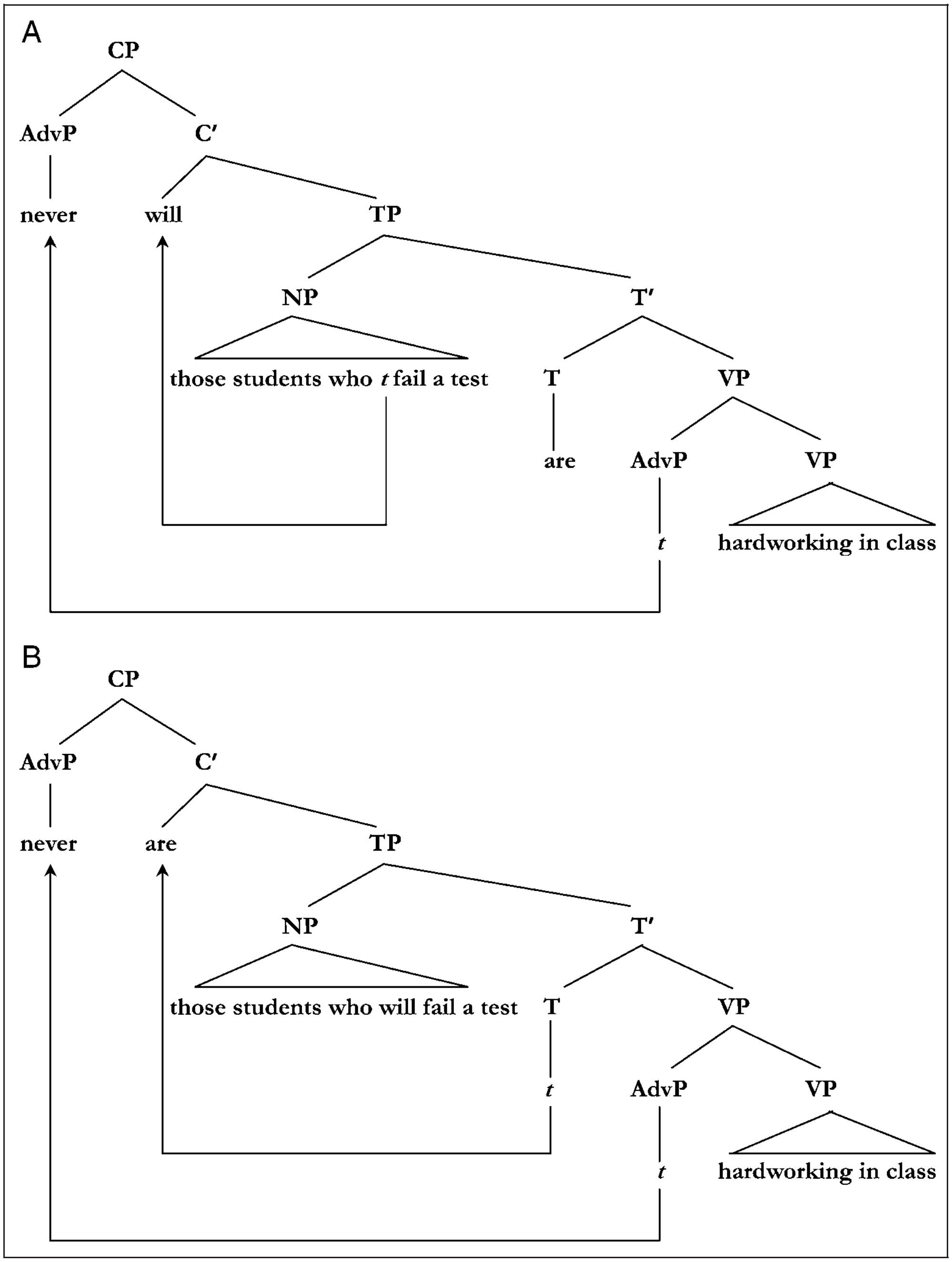

**Figure 2.** The tree structures consist of hierarchically defined relations representing how sentences are constructed. The detailed tree structures are irrelevant to the present discussion, but *t* indicates the base position of a moved element. Of the three auxiliaries (*will*, *have*, and *are*), it is always the highest auxiliary in terms of the tree structure that is fronted to make grammatical NI sentences (B) and (C) rather than the first auxiliary in terms of the linear order as in (A) and (D). CP = complementizer phrase; TP = tense phrase; VP = verb phrase; AdvP = adverb phrase.

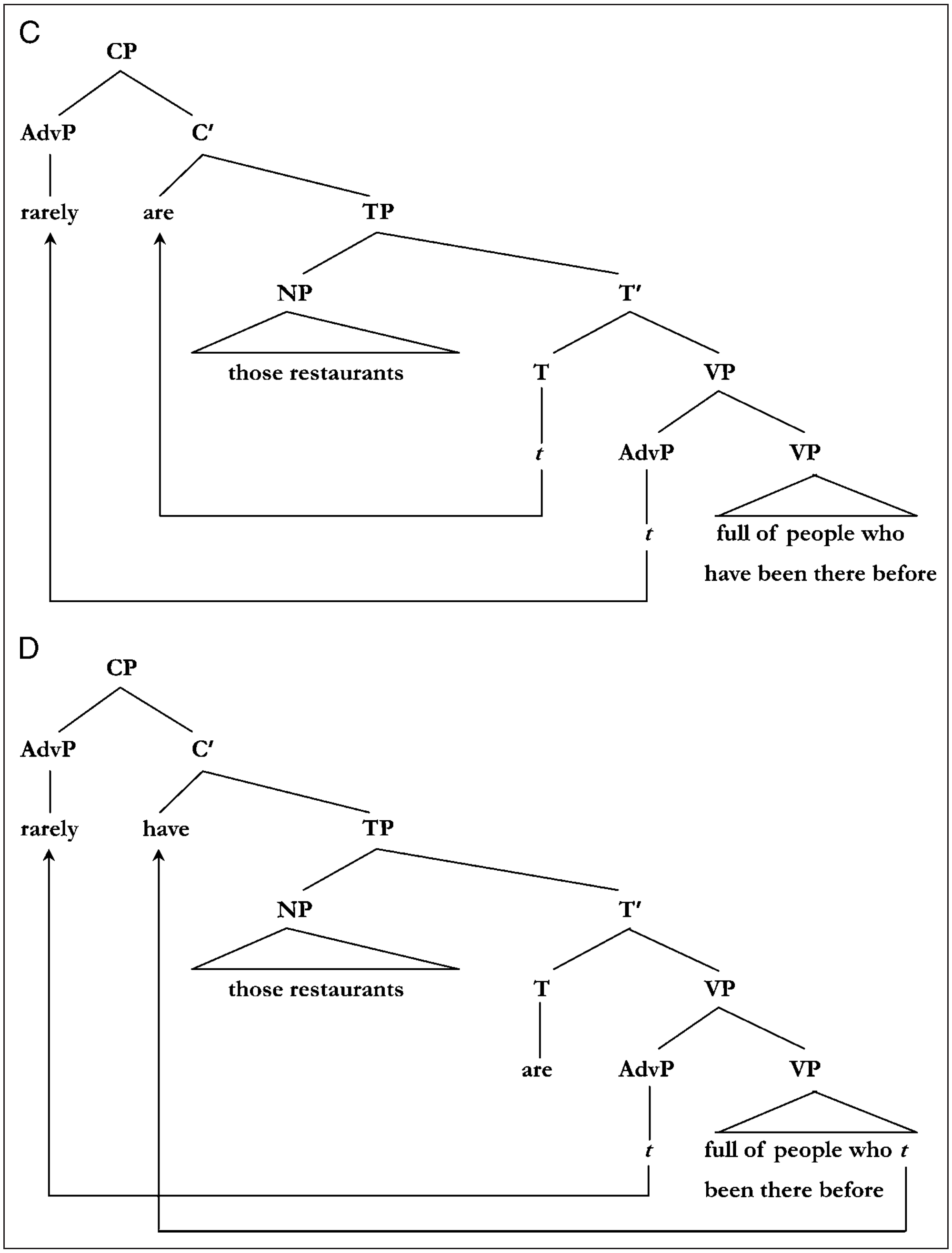

**Figure 2.** (*continued*).

independent rules that just operate on the sequence of words (Chomsky, 1980). The principle of structure dependence rules out impossible rules that are blind to hierarchical structure.

Previous behavioral studies tested the principle of structure dependence with simpler examples of question sentence formation than NI, reaching the conclusion that the principle cannot be inferred from the input and, therefore, forms part of the human language recipe: English-speaking children aged 3:2 to 5:10 do not produce yes/no questions violating the principle of structure dependence in their mother tongue (Crain & Nakayama, 1987); even an unprecedented polyglot savant endowed with the facility of picking up over 20 languages cannot learn an invented language that has a number of impossible, structure-independent rules (Smith & Tsimpli, 1996). Thus, assuming that structure dependence is a universal principle applying to every human language including L2, our hypothesis is that L2 learners should respect a structure-dependent rule as L1 speakers do. In addition, recent neuroimaging studies provide convergent evidence that L1 and L2 are processed in the same cortical regions, suggesting the possibility that L1 and L2 share the same neural devices (Perani & Abutalebi, 2005; Tatsuno & Sakai, 2005; Sakai, Miura, Narafu, & Muraishi, 2004; Musso et al., 2003; Tettamanti, Alkadhi, Moro, &

Perani, 2002). Specifically, syntactic rules respecting the principle of structure dependence selectively activate Broca's area, including the pars triangularis (Musso et al., 2003) and the pars opercularis (Tettamanti et al., 2002).

## METHODS

In the present study, two groups were tested: The instruction and non-instruction groups both consisted of 20 adult Japanese learners of English in the same department at a university in Japan. Therefore, both groups were exposed daily to English in English classes at the university. Both groups were indistinguishable from each other in terms of their overall English proficiency estimated by the Test of English for International Communication (TOEIC, a standardized test) and knowledge of NI in simplex sentences at the time of the first functional magnetic resonance imaging (fMRI) measurement (Test 1). At the first fMRI, neither group was assumed to possess knowledge of NI, judging from the fact that they could not correctly distinguish the grammatical sentences such as *Never are those students late for class* from the ungrammatical sentences such as *Never those students are late for class*. The instruction group received instruction for NI only with *simplex* sentences (e.g., *Those students are never late for class → Never are those students late for class*) after Test 1. Participants in the instruction group met twice a week for one month (8 classes in total), with one training session lasting an hour in addition to their regular classes, and were required to hand in assignments based on the training sessions. More crucially, no instruction was given to the instruction group with regard to *complex* sentences containing relative clauses (e.g., *Those students who are very smart are never silent in class*), which the structure-dependent rule can deal with but the structure-independent one cannot. In contrast, the non-instruction group received no instruction whatsoever after Test 1. Note here that participants in the non-instruction group were exposed to English in their regular classes at the university as were those in the instruction group. Therefore, the difference between the two groups was the presence/absence of the instruction of NI with simplex sentences in the training sessions.

Pre- and post-instruction tests were administered to determine which rule (structure-dependent or structure-independent) our participants would use with complex sentences as well as simple sentences. The instruction group participated in instruction sessions for one month in an EFL context (i.e., classroom-based L2 instruction in Japan) and underwent MRI scannings before the instruction (Test 1) and after the instruction (Test 2). As with the instruction group, the non-instruction group also underwent two fMRI measurements. In both Test 1 and Test 2, MRI scanning consisted of two sessions (Session 1 and Session 2), where the participants viewed a series of three or four sequentially presented stimuli on the screen—three phrases in Session 1 and four phrases in Session 2—in a randomized order: Note here that the term "phrase" in this article is meant to refer to a block of words or a fixation cross (i.e., +) (Table 2). In Test 2, the participants were

**Table 2.** Sample Sentences Used in the fMRI Experiment

| | | | | |
|---|---|---|---|---|
| *Session 1: 140 Simplex Sentences (Instruction Only for the Experimental Group)* | | | | |
| GC-s | Those students are never | late for class: | + | |
| UC-s | Those students are | late for never class: | + | |
| GI-s | Never are those students | late for class: | + | |
| UI-s | Never those students are | late for class: | + | |
| | 1700 msec | 1700 msec | 2000 msec | |
| *Session 2: 140 Complex Sentences (No Instruction for Either Group)* | | | | |
| GC-c | Those students | who are very smart | are never silent in class: | + |
| UC-c | Those never students | who are very smart | are silent in class: | + |
| GI-c | Never are those students | who are very smart | silent in class: | + |
| UI-c | Never are those students | who very smart | are silent in class: | + |
| | 1700 msec | 1700 msec | 1700 msec | 2000 msec |

The experiment was performed in an event-related design. All stimuli were randomly presented onto the screen in an automatic phrase-by-phrase manner. Time length of one phrase was 1700 msec. Participants were asked to judge whether the sentences they saw on the screen were syntactically correct, except for the N condition. Response time was recorded from the onset of the second phrase in Session 1 or the third phrase in Session 2 to the button press for each stimulus (maximum = 3700 msec).

GC condition = grammatical canonical order condition; UC condition = ungrammatical canonical order condition; GI condition = grammatical inversion condition; UI condition = ungrammatical inversion condition; -s = simplex sentences; -c = complex sentences.

tested on a set of sentences different from those used in Test 1 and in the instruction sessions so as to exclude the effects of rote learning and memorization.

Session 1 comprised two types of simplex sentences as well as a null condition: canonical order and inversion, grammatical and ungrammatical. Hence, there were five conditions: grammatical canonical order simplex (GC-s) sentences, ungrammatical canonical order complex (UC-s) sentences, grammatical inversion simplex (GI-s) sentences, ungrammatical inversion simplex (UI-s) sentences, and a null condition (N; i.e., a fixation cross on the screen). Typical stimuli of the conditions are given in Table 2. The GC-s condition tested the fundamental knowledge of the canonical position of negative adverbs; the UC-s condition tested the knowledge of the incorrect position of negative adverbs; the GI-s condition tested the knowledge of NI involving the negative inversion of auxiliaries; the UI-s condition tested the knowledge of ungrammatical NI sentences, which lack NI. Session 2 consisted of complex sentences with relative clauses (i.e., GC-c, UC-c, GI-c and UI-c sentences) as well as a null (N) condition. The GC-c and UC-c conditions tested the knowledge of the grammatical and ungrammatical positions of negative adverbs in complex sentences, respectively. GI-c sentences (i.e., grammatical inversion complex sentences) are derived from their respective canonical order sentences in accordance with the principle of structure dependence, whereas UI-c sentences (i.e., ungrammatical inversion complex sentences) violate the principle. The tasks differed in the instructions informing the participants on how to respond to the stimuli: (1) perform the grammaticality identification task (judging whether the sentences were grammatically correct or not) by clicking one of two buttons in the second phrase in Session 1 and the third phrase in Session 2; a colon (:) put after each stimulus indicates the phrase where participants were required to judge grammaticality of the stimulus; (2) passively view a fixation cross on the null condition.

If structure dependence is a universal principle that is biologically constrained and hardwired in our brain, and is still operative in L2 acquisition, then adult Japanese learners of English in the instruction group should make exclusive use of the principle and succeed in comprehending uninstructed complex NI sentences on the basis of their exposure to simplex NI sentences in the training sessions. In contrast, those in the non-instruction group might not improve their understanding of complex NI sentences because they still lack the knowledge of simplex NI sentences, which will not enable them to know more than they have experienced.

## Participants

Forty Japanese native speakers participated in the experiment and were randomly divided into instruction and non-instruction groups on the basis of their TOEIC scores and their error rates for NI in simplex sentences at the first fMRI experiment. However, only 17 participants in the instruction (INS) group (mean ages ± *SD*: 21.6 ± 1.6 years) and 19 participants in the non-instruction (NON-INS) group (mean ages ± *SD*: 22.0 ± 2.3 years) were included in the data analysis. The exclusion of the 4 among the total of 40 participants was due to technical problems. There were no significant differences between the two groups in their mean scores on the TOEIC and error rates for the first fMRI scan (INS: $r = -.24$, $p = .35$, NON-INS: $r = -.29$, $p = .23$). Therefore, the two groups were considered to be qualitatively comparable in English knowledge at the time of the first fMRI measurement. All of the participants began to study English as a second language in Japanese schools (the age of first exposure to English in a formal school setting in the instruction group was 12.4 ± 0.3 years and the mean age in the non-instruction group was 12.5 ± 0.3 years).

All participants gave written informed consent for the study and right-handedness was verified using the Edinburgh Inventory (Oldfield, 1971). All experiments were performed in compliance with the relevant institutional guidelines approved by Tohoku University. Before the fMRI experiments, each participant received a list of English words or phrases with their translations to be used in the experiments and was requested to learn the words so that they would have no difficulty with semantics of stimulus sentences used in the fMRI measurements. After each fMRI scanning, we tested the knowledge of words in the list. To familiarize participants with tasks inside an fMRI scanner, a trial sequence for sample stimuli was administered with a computer before each fMRI scanning.

## Stimuli and Tasks

For the fMRI measurements, we selected 140 simplex sentences for Session 1 and 140 complex sentences for Session 2 (Table 2). The sentences were presented to participants in an event-related design with Session 1 followed by a break outside the fMRI scanner and then Session 2: Each session contained five condition tasks, GC (grammatical canonical order sentence), UC (ungrammatical canonical order sentence), GI (grammatical inversion sentence), UI (ungrammatical inversion sentence), and N (null task). Each condition consisted of 35 sentences and fixation crosses. The sentences in Sessions 1 and 2 were divided into three and four phrases, respectively, and were presented phrase by phrase visually to the participants lying in the scanner. Each phrase was shown on the screen for 1.7 sec, followed by a fixation cross for 2 sec (Table 2). In each condition, the participants were instructed to judge the grammaticality of the stimuli after entering the second phrase in Session 1 and the third phrase in Session 2, except for the N condition, in which they were advised to gaze at a fixation cross. All stimuli were controlled using E-prime (Psychology Software Tools).

## fMRI Data Acquisition and Analyses

Data recordings were obtained with fMRI using a 1.5-Tesla Siemens Symphony Scanner (Siemens, Erlangen, Germany) at Tohoku University, under the preliminarily determined condition in echo-planar imaging (TR = 2300 msec; TE = 50 msec; slices = 25; thickness = 5 mm; flip angle = 90°; FoV = 192 mm). After the attainment of functional imaging, anatomical T1-weighted MDEFT images (thickness = 1 mm; FoV = 256 mm; data matrix = 192 × 224; TR = 1900 msec; TE = 3.93 msec) were also acquired from all the participants. Excluding the 12 dummy scans for stabilization of the T1-saturation effect, we acquired 417 (Session 1) and 513 (Session 2) scans for each subject.

To specify the cortical region exhibiting activation during the various tasks, statistical analyses of image processing were carried out using SPM2, operating on a MatLab platform, developed by the Wellcome Department of Cognitive Neurology, University of London, UK. The effects of head motion across scans were corrected by realigning all the scans to the first one. The realigned datasets were spatially normalized to the Montreal Neurological Institute standard brain template, and then we performed smoothing of images using a 10-mm Gaussian filter. An analysis of the tasks for each participant was conducted at the first statistical stage and a group statistical analysis was performed at the second stage. We set the statistical threshold at $p < .05$ (corrected for family-wise error rate [FWE]). Finally, we performed region-of-interest (ROI) analyses in brain areas obtained from the GI-c–N(Test 2 − Test 1) and UI-c–N(Test 2 − Test 1) analyses. The locations of cortical activation specified by the interconditional subtraction were closely mapped out on the basis of the *Co-Planar Stereotaxic Atlas of the Human Brain* (Talairach & Tournoux, 1988).

## RESULTS

### Behavioral Results

We analyzed data from 17 participants in the instruction group and 19 participants in the non-instruction group. Both groups had knowledge of grammatical canonical word order (GC) of negative adverbs in simplex sentences at the time of Test 1 [% of error rates: INS, 13.4 ± 13.3 (mean ± *SD*); NON-INS, 10.8 ± 13.1], but neither had knowledge of NI with simplex sentences (GI-s: INS, 42.7 ± 35.4; NON-INS, 37.3 ± 32.5; UI-s: INS, 46.1 ± 37.0; NON-INS, 40.5 ± 28.5) (Figure 3A1, A2), let alone complex sentences (GI-c: INS, 40.7 ± 37.3; NON-INS, 35.9 ± 26.5; UI-c: INS, 50.3 ± 34.6; NON-INS, 62.3 ± 26.7) (Figure 3B1, B2). There was no significant difference in the percentage of error rates in Session 1 [GI-s: $t(34) = 0.48$, *ns*; UI-s: $t(34) = 0.51$, *ns*] and Session 2 [GI-c: $t(34) = 0.44$, *ns*; UI-c: $t(34) = -1.16$, *ns*] in Test 1 between the instruction group and the non-instruction group. These results indicated that both groups had developed the same level

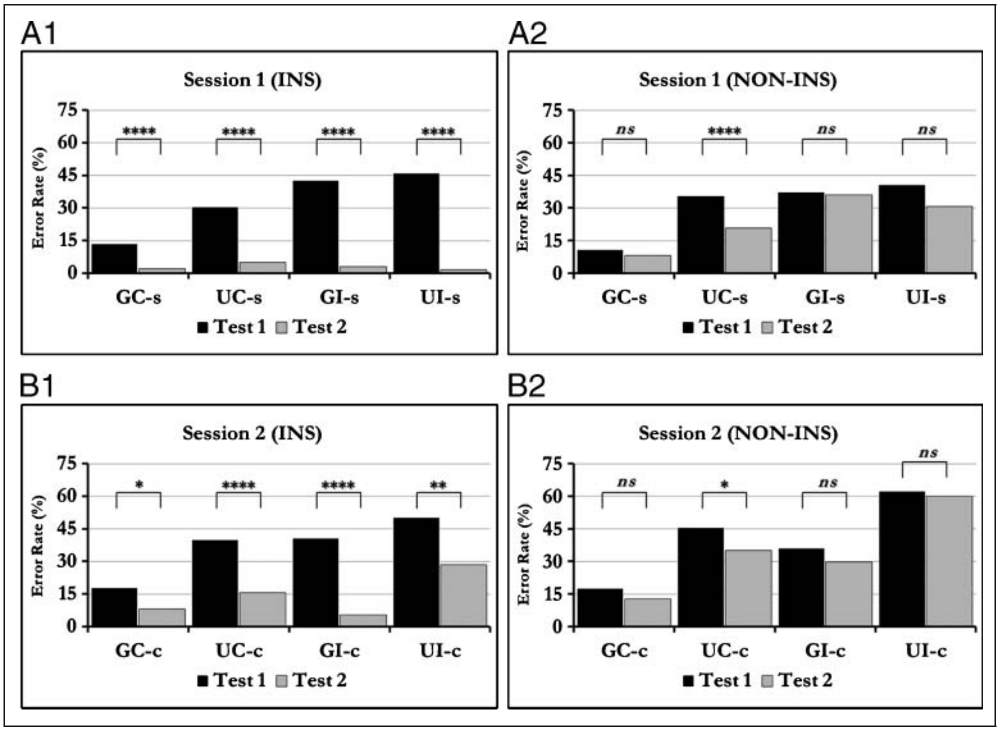

**Figure 3.** Behavioral data for the instruction group (INS) and the non-instruction group (NON-INS): error rate (%) for the two groups in judging the syntactical correctness of the sentences presented during the two sessions. Session 1 consisted of 140 simplex sentences for which the INS had received instruction, whereas Session 2 consisted of 140 complex sentences involving relative clauses for which neither the INS nor the NON-INS had received instruction. The INS clearly showed a significant improvement in all task conditions from Test 1 (pre-instruction) to Test 2 (post-instruction) (A1 and B1), whereas the NON-INS remained almost the same across the two tests, except in the UC condition (A2 and B2). This improvement in the UC condition for the NON-INS could possibly be due to "a familiarization effect." GC = grammatical canonical order condition; UC = ungrammatical canonical order condition; GI = grammatical inversion condition; UI = ungrammatical inversion condition; INS = instruction group; NON-INS = non-instruction group. *$p < .05$, **$p < .01$, ****$p < .001$, *ns* = not significant.

of knowledge of the NI structure in English at the time of Test 1.

The percentage of errors in Session 1 of Test 2, however, significantly decreased with the instruction group, suggesting that the participants in the instruction group acquired knowledge of NI after one month of instruction [GC-s: $t(16) = 4.17$, $p < .001$; UC-s: $t(16) = 5.35$, $p < .001$; GI-s: $t(16) = 4.71$, $p < .001$; UI-s: $t(16) = 4.85$, $p < .001$] (Figure 3A1). Most intriguing and surprising is the finding that the participants in the instruction group acquired knowledge of NI with complex sentences for which they had *never* received instruction [GC-c: $t(16) = 2.60$, $p < .05$; UC-c: $t(16) = 5.53$, $p < .001$; GI-c: $t(16) = 3.93$, $p < .001$; UI-c: $t(16) = 3.10$, $p < .01$] (Figure 3B1): They overwhelmingly rejected the complex sentences violating structure dependence (UI-c: 28.6 ± 21.4), while correctly accepting those without the violation of structure dependence (GI-c: 5.4 ± 5.8). The results in Test 2 show that participants in the instruction group developed a mental grammar which not only encompassed the simplex sentences for which they had received instruction but also transferred this knowledge to aid in the comprehension of complex sentences for which there had been no instruction (Figure 3A1 and B1). In contrast, as expected, no significant improvement in accuracy was observed for the non-instruction group in Test 2, except for the UC condition [GC-s: $t(18) = 0.95$, *ns*; UC-s: $t(18) = 4.46$, $p < .001$; GI-s: $t(18) = 0.28$, *ns*; UI-s: $t(18) = 1.89$, *ns*; GC-c: $t(18) = 1.01$, *ns*; UC-c: $t(18) = 2.32$, $p < .05$; GI-c: $t(18) = 1.81$, *ns*; UI-c: $t(18) = 0.60$, *ns*] (Figure 3A2 and B2). The improvement on the UC-s and UC-c conditions in Test 2 for the non-instruction group could possibly be due to a familiarization effect on the canonical word order of negative adverbs in non-NI sentences: The non-instruction group might have been familiarized with ungrammatical canonical simplex sentences (UC-s) and complex sentences (UC-c) which they had not been exposed to until Test 1, probably because the ungrammatical position of negative adverbs in noninversion simplex sentences (UC-s) and complex sentences (UC-c) was easy to detect on the basis of their knowledge of GC-s, and exposure to UC-s in Test 1 made the participants in the non-instruction group sensitive to the ungrammaticality of UC-s and UC-c. The ungrammatical canonical sentences (UC-s and UC-c) in Table 2 show that the ungrammaticality of UC-s and UC-c can easily be probed only by checking what word (s) the negative adverb "never" is merged with in a linear order. The lack of a familiarization effect with regard to the GC-s and GC-c conditions for the non-instruction group could also be explained by the fact that the judgments of grammatical canonical sentences require more comprehensive knowledge of English syntax than those of ungrammatical canonical sentences: In order to judge whether a canonical sentence is ungrammatical, only a piece of information regarding the wrong concatenation in a local phrase (e.g., "*never* class" in UC-s in Table 2) is sufficient, even though the participants read the whole sentence; in order to judge whether a canonical sentence is grammatical, the whole structure, that is, every concatenation of words, must be checked, deferring a familiarization effect on the basis of the knowledge of GC-s and exposure to stimuli in Test 1. It is more essential to keep in mind here that NI sentences (GI-s and UI-s) for the non-instruction group did not cause a familiarization effect as sensitivity to the NI sentences did require syntactic knowledge of NI, which was not inferred from their knowledge of GC-s.

## Imaging Data

In order to identify cortical activation generated after the instruction, we subtracted $(\mathrm{GI} - \mathrm{N})_{\mathrm{TEST1}}$ from $(\mathrm{GI} - \mathrm{N})_{\mathrm{TEST2}}$, and $(\mathrm{UI} - \mathrm{N})_{\mathrm{TEST1}}$ from $(\mathrm{UI} - \mathrm{N})_{\mathrm{TEST2}}$, where N stands for the baseline (N) task (looking at a fixation cross on the screen). Concerning the fMRI data on the inversion conditions (GI-c, UI-c) in complex sentences (Session 2), significant activation was observed only for the *instruction* group: the left inferior frontal gyrus (IFG) and the bilateral precuneus in the GI-c − N condition; the bilateral IFG, the bilateral precuneus, the right middle frontal gyrus (MFG), and the right superior parietal lobule (SPL) in the UI-c − N condition (Table 3).

An ROI analysis was then conducted in the left IFG (GI-c − N: −48, 18, 18; UI-c − N: −46, 16, 24), which is assumed to be involved in the processing of language. There was no significant difference between other conditions in the left IFG. No significant signal change in GI-s − N in Session 1 (simplex sentences) between Test 1 and Test 2 was found for the instruction and non-instruction groups (Figure 4A1). Regarding the signal change in UI-s in Session 1 between Test 1 and Test 2, the instruction group showed significantly reduced activations in the left IFG, whereas the non-instruction group showed no significant signal change (Figure 4B1). Notably, significant activations in the left IFG were found for the uninstructed complex sentences in Session 2 after the instruction (Figure 4A2). More specifically, the pars triangularis [F3t or Brodmann's area (BA 44/45)] of the left IFG was more significantly activated during the measurement of inversion conditions (GI-c − N and UI-c − N) in Session 2 in Test 2 than in Test 1 (Figure 4A2 and B2). A left IFG activation decrease in the UI-s − N condition in Test 2 for the instruction group (Figure 4B1) may imply that one-month instruction of NI using simplex sentences had already been translated into the acquisition and consolidation of the knowledge of NI. This is consistent with the view that an increasing cortical activation during the process of learning a new rule is followed by the maintenance of the increased activation and then a decreased activation as a result of the consolidation of the rule (Indefrey, 2006; Tatsuno & Sakai, 2005; Sakai et al., 2004). In contrast, no significant cortical activation change in Session 1 and

**Table 3.** Activated Regions in the Contrast of GI-c − N(Test 2 − Test 1) and UI-c − N(Test 2 − Test 1)

| *Brain Regions* | *BA* | *Side* | $t_{max}$ ($Z_{max}$) | *x* | *y* | *z* | *Cluster Size* |
|---|---|---|---|---|---|---|---|
| *GI-c − N (Test 2 − Test 1)* | | | | | | | |
| Precuneus | 7 | R | 13.32 (6.24) | 14 | −76 | 48 | 183 |
| | 7 | L | 9.96 (5.55) | −12 | −74 | 50 | 260 |
| | 7 | L | 9.46 (5.42) | −20 | −76 | 52 | 260 |
| | 19 | L | 7.83 (4.95) | −30 | −76 | 38 | 260 |
| | 19 | R | 8.42 (5.13) | 32 | −64 | 40 | 133 |
| | 19 | R | 8.20 (5.07) | 34 | −72 | 38 | 133 |
| IFG | 44/45 | L | 11.01 (5.79) | −48 | 18 | 18 | 23 |
| *UI-c − N (Test 2 − Test 1)* | | | | | | | |
| Precuneus | 19 | R | 14.98 (6.50) | 34 | −66 | 38 | 453 |
| | 19 | L | 11.60 (5.91) | −28 | −74 | 38 | 105 |
| | 7 | L | 11.07 (5.80) | −10 | −74 | 46 | 249 |
| | 7 | R | 9.93 (5.54) | 16 | −72 | 46 | 177 |
| MFG | 6/8/9 | R | 10.53 (5.68) | 52 | 8 | 38 | 198 |
| | 9 | R | 7.82 (4.95) | 48 | 20 | 30 | 198 |
| SPL | 7 | L | 8.59 (5.18) | −4 | −68 | 55 | 249 |
| | 7/40 | R | 10.16 (5.60) | 38 | −54 | 48 | 453 |
| IFG | 6/9 | R | 9.57 (5.45) | 46 | 4 | 32 | 198 |
| | 9/44 | L | 9.08 (5.32) | −46 | 16 | 24 | 20 |

Respective activated anatomic region, approximate Brodmann's area, right or left (R, L) hemisphere, $t$ values, and $Z$ values. Stereotactic coordinates ($x, y, z$) as defined by MNI are shown for each voxel with a local maximum of $t$ values in the contrasts indicated ($p < .05$, corrected). IFG = inferior frontal gyrus; MFG = middle frontal gyrus; SPL = superior parietal lobule.

Session 2 between Test 1 and Test 2 was detected for the non-instruction group (Figure 4).

## DISCUSSION

The activations in the precuneus are not within the scope of the present article, but they are likely to reflect visuospatial attention task demands in our experiments (Cavanna & Trimble, 2006), but an fMRI study of the acquisition of structure-dependent and structure-independent rules with the visual representation of sentences also reveals the activation of the precuneus in addition to Broca's area: Visuospatial processing might be involved in generating "detailed representations of ordered sequences of symbols, such as words forming written sentences" (Tettamanti et al., 2002). The precuneus is not traditionally viewed as being involved in language processing, but there is increasing evidence that the precuneus supports syntax processing: The precuneus is engaged in the processing of hierarchical structure generally and syntactic structure in particular (Bachrach, 2008), which is consistent with the result in our experiment (see also Caplan, Waters, & Alpert, 2003, for manipulation of syntactic complexity in the precuneus and Shetreet, Friedmann, & Hadar, 2009 for the evidence that the precuneus is more language-specific than has previously been assumed).

The cortical activation changes in the left IFG between Test 1 and Test 2 are worth mentioning. No significant change was found with regard to the GI-s − N condition (grammatical simplex sentences) for the instruction and non-instruction groups. Judging from the percentage of errors in GI-s, the instruction group acquired knowledge of NI with an almost-perfect performance after one month of instruction, whereas the non-instruction group did not learn NI (Figure 3A1 and A2). The lack of the signal change in the GI-s − N condition for the non-instruction group is trivial, judging from the lack of the improvement in error rates. A possible interpretation for the lack of the cortical activation change despite the significant decrease of the errors in the instruction group would be that the participants in the instruction group must have shown the initial increase in activations in earlier stages of learning and at the time of Test 2 they had already consolidated the

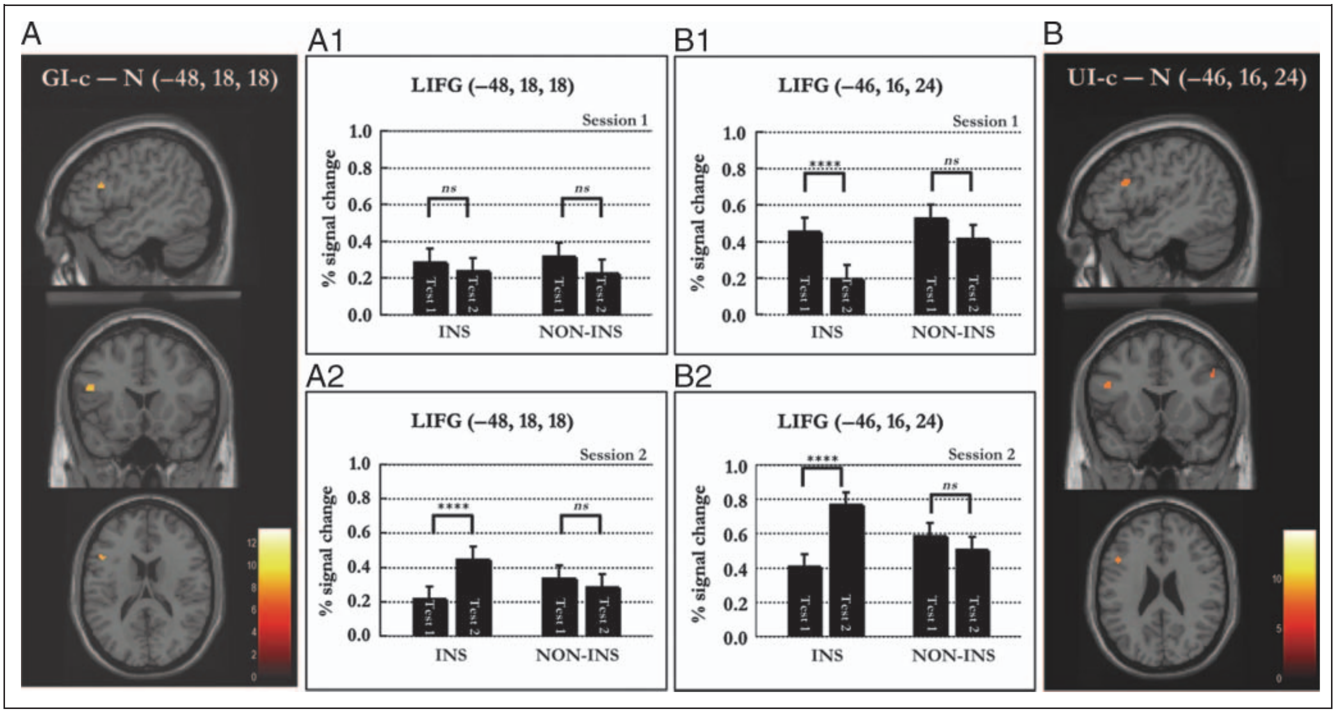

**Figure 4.** Brain activation and ROI analysis for the pars triangularis of the left IFG (LIFG). (A1) No significant signal change in GI-s − N in Session 1 (simplex sentences) between Test 1 and Test 2 for either group. (B1) An LIFG activation decrease in the UI-s − N condition in Session 1 (simplex sentences) in Test 2 for the INS might imply that one-month instruction of NI using simplex sentences had already been translated into the acquisition and consolidation of the knowledge of NI. (A2 and B2) Significant activations in the LIFG were found for the uninstructed complex sentences in Session 2 (complex sentences) after the instruction for the INS. More specifically, the pars triangularis of the LIFG was more significantly activated during the measurement of inversion conditions (GI-c − N and UI-c − N) in Session 2 in Test 2 than in Test 1. In contrast, no significant cortical activation change in Session 1 and Session 2 between Test 1 and Test 2 was detected for the non-instruction group (NON-INS). ****$p < .001$, *ns* = not significant.

knowledge of NI due to repeated practice of simplex NI sentences, showing "seemingly no activation change." Regarding the UI-s − N condition, the instruction group showed significantly weaker activations in the left IFG in Test 2 due to the consolidation of the knowledge of NI. Note here that there was no significant difference in cortical signal changes in Test 2 between the GI-s − N and UI-s − N conditions for the instruction group [$t(16) = 1.06, p = .307$] (Figure 4A1 and B1).

Our results showed a significant activation in the left IFG (F3t) that is revealed by the overall comparison of post-instruction (Test 2) minus pre-instruction (Test 1) in the processing of "uninstructed complex NI sentences" (Session 2) (Figure 4A and A2 for the GI-c − N condition and Figure 4B and B2 for the UI-c− N condition). The present study thus provides not only the behavioral evidence (GI-c and UI-c in Figure 3B1) but also the first neuroimaging evidence (GI-c − N in Figure 4A2 and UI-c − N in Figure 4B2) to clearly demonstrate that L2 learners know more than they could have learned from the data they have encountered, suggesting that there is much more to L2 acquisition than merely learning what is instructed, and pointing to the interweaving of nature (UG) and nurture (instruction) in L2 acquisition. In this respect, this study provides new insight into the role of Broca's area, where nature (UG) and nurture (instruction) contribute to enabling L2 learners to know more than is instructed. A crucial point here is in the linkage between the acquisition of new structures that were not instructed and the accompanying cortical signal changes in the left IFG.

Our result fits well with recent imaging studies in that F3t subserves the acquisition of a real syntactic rule (Tatsuno & Sakai, 2005; Sakai et al., 2004; Musso et al., 2003), and that Broca's area (BA 44/45) supports hierarchical dependencies and long-distance dependencies (Friederici, Bahlmann, Heim, Schubotz, & Anwander, 2006; Ben-Shachar, Palti, & Grodzinsky, 2004; Tettamanti et al., 2002; for the different location of activation in Broca's area between BA 44 and BA 45, see Makuuchi, Bahlmann, Anwander, & Friederici, 2009; Bahlmann, Schubotz, & Friederici, 2008). F3t is reported to play a role in processing sentences (for a meta-analysis, see Vigneau et al., 2006) and sentences with high demands on memory (Makuuchi et al., 2009). Activation in F3t in the present study is all the more interesting because NI is vanishingly rare in the usage of English-speaking people and feels literary, old-fashioned, and sophisticated (Slobin, 2003) but it is subject to the principle of structure dependence. Knowledge of a new rule (NI) during the instruction session using simplex sentences, amplified by the universal principle of structure dependence, is conjectured to have projected into a rich knowledge of complex NI sentences that L2 learners have not been taught. Put another way, the ontogenesis of L2 knowledge about NI in complex sentences was driven by the interaction of the UG principle (i.e., structure dependence) and the linguistic experience (the learning of NI in simplex sentences during the training session). Thus, this study suggests that the brain remains plastic in adults L2 learners and that there is no critical period in L2 acquisition, at least in the domain of core principles of UG (i.e., the principle of structure dependence) that characterize possible human languages (see Birdsong, 2006 for an insightful overview and critiques of a critical period hypothesis; Kotz, 2009 for a review of the findings of the L2 literature on the neural correlates and mechanisms of syntax acquisition that do not support a critical period hypothesis in L2 acquisition).

It might be objected that "long-distance" dependencies are crucially involved in the grammaticality judgments of complex NI sentences, so that it is difficult, if not impossible, to tease the acquisition of the new rule NI apart from effects of long-distance dependencies. This objection is, however, based on long-distance linear dependencies. Note, incidentally, that GI-c and UI-c sentences in our stimuli exhibit short-distance linear dependencies as well as long-distance linear dependencies (e.g., (2) and (3) in Table 1), thus it is implausible to assume that activations of the left IFG were totally due to long-distance linear dependencies. If effects of long-distance linear dependencies had exclusively been responsible for the activations of the left IFG, then the same cortical signal change should have been observed for the non-instruction group. Structurally speaking, however, movement of the matrix auxiliary in complex NI sentences does not show long-distance hierarchical dependencies because it crosses one sentence boundary (e.g., the matrix TP in Figure 2B and C). On the other hand, movement of the auxiliary embedded in a relative clause (e.g., *will* in Figure 2A and D) exhibits a long-distance hierarchical dependency because it crosses two sentence boundaries (i.e., the embedded relative clause in addition to the matrix TP in Figure 2A). It is always the highest auxiliary in terms of the hierarchical structure rather than the first auxiliary in terms of the linear order that is fronted to make grammatical NI sentences, simplex or complex. In terms of hierarchical dependencies, the auxiliaries that legitimately move to generate GI sentences are "structurally closer" to the fronted negative adverbs than those which illegitimately move to generate UI sentences. In fact, GI sentences, simplex or complex, always exhibit short-distance hierarchical dependencies, despite their apparent long-distance linear dependencies in some cases. Remember that the principle of structure dependence refers to a hierarchical position rather than a linear position. Therefore, long-distance linear dependencies are not relevant to the operation of NI.

Several studies have investigated the brain activation of a new rule after instruction but they used the *same* sentence or verb patterns used during instruction sessions (Sakai et al., 2004; Musso et al., 2003; Tettamanti et al., 2002). As a consequence, they cannot completely eliminate the possibility that the participants in those studies might have just memorized the relevant rules and then repeated back what they had learned (Marcus, Vouloumanos, & Sag, 2003). However, our study succeeded in overcoming this limitation by using complex sentences involving relative clauses in Test 2, which were structurally different from simplex sentences used during the instruction sessions. Nevertheless, it is striking that activation was observed in the same area (F3t) that previous studies had identified as being responsible for the acquisition of a new rule (Musso et al., 2003). This finding cannot possibly be accommodated within the assertion that L2 learners are not creative but simply imitate what they are instructed. This study does not claim that imitation is irrelevant for L2 acquisition, but it does not suffice to tell the whole story of L2 acquisition. It is also implausible that L2 learners have utilized general problem-solving abilities. The difference in activation of F3t in Test 2 between the instruction and non-instruction groups cannot be explained either in terms of frequency effects, working memory, or syntactic complexity (Marcus et al., 2003). Moreover, it is crucial that exposure to the relevant L2 input in this study was not brief for the instruction group as in a laboratory setting of a previous study (Opitz & Friederici, 2004; Musso et al., 2003), providing the participants with the situations where linguistic knowledge was engaged. The results also provide more compelling evidence for the core computational mechanism of L2 learners than previous studies using an artificial grammar learning paradigm (Friederici et al., 2006; Opitz & Friederici, 2004), in that the complex NI sentences in our study involve "nesting relations" that cannot be processed by a counting strategy plus memory. The counting mechanism can deal with the center-embedded hierarchical structures ($A^nB^n$) generated by hierarchical phrase structure grammar, where A and B represent symbols and $n$ the number of occurrences, that is, $n$ occurrences of A followed by $n$ occurrences of B (see Bahlmann et al., 2008 for the artificial grammar research removing the possibility of employing a counting strategy). Incidentally, note that a success with the $A^nB^n$ grammar reported in the artificial paradigm mentioned above as well as the experiment on starlings (Gentner, Fenn, Margoliash, & Nusbaum, 2006) does not indicate the acquisition of the computational system that generates an infinite array of the hierarchical structures manifest in human languages (Fitch, 2010; Hauser, 2010) (for a critical review of Friederici et al., 2006, see Piattelli-Palmarini, Uriagereka, & Salaburu, 2009).

The result here also demonstrates that age of L2 acquisition does not affect the processing of a core (unparameterized) syntactic principle. Yet, this stands in direct contrast to a number of L2 studies on critical or sensitive periods demonstrating that age of L2 acquisition is crucial for the acquisition of syntax (Wartenburger et al., 2003; DeKeyser, 2000; Johnson & Newport, 1989).

The difficulties reported in those studies, however, mostly pertain to L2 knowledge of the agreement marking of number, gender, or case on articles, nouns, or verbs. It is important to note that these linguistic phenomena exhibit parametric variations across languages (i.e., presence or absence of overt markings of articles or verbs). The agreement obeys a universal syntactic constraint (e.g., locality), but the existence of overt morphology is not universal (Hale, 1996). For example, English exhibits an overt agreement marking on verbs, but Japanese does not. It is worth noting that linguistic knowledge or mental grammar comprises modules such as lexicon, syntax, morphology, phonology, semantics, pragmatics, and so forth, as well as their coordinations, interfaces. Recent research on "fossilized" L2 knowledge uncovers the fact that even advanced L2 learners have permanent difficulty with inflections marking number, tense, case, gender, as well as function words such as articles (Sorace, 2005; Hawkins, 2000), probably because those phenomena are not purely syntactic but are related to the interface between syntax and other domains (phonology, lexicon, discourse/pragmatics). In other words, the syntactic stimuli used in previous studies to argue for the age effects on the acquisition of L2 syntax could be interpreted as being purely nonsyntactic, that is, being external to the computational system of generating syntactic structures.

More importantly, knowledge of structure dependence was investigated in a few neuroimaging studies (Musso et al., 2003; Tettamanti et al., 2002). It should be noted that structure dependence is an abstract principle without any overt realization in any language, making it difficult for learners (L1 or L2) to receive enough information of structure dependence. Because it is one of the fundamental unparameterized principles that apply to every natural language, L2 learners never fail to make use of it through their lives, given that L2 is a natural language. As far as the principle of structure dependence is concerned, it is impossible for L1 and L2 acquisition to be fundamentally different. The initial state of the innate language faculty characterized by UG develops into stable states characterized by a particular adult grammar such as English grammar or Japanese grammar, depending on the kinds of linguistic input to which we are exposed (Chomsky, 2004a). In this view, it is no wonder that "to have knowledge of L1 is necessarily to have knowledge of UG (Hale, 1996)." In other words, (some) UG principles have already been activated when adults have acquired a native language. Assuming the operation of structure dependence (UG principle) in some or another form in every human language, it appears to be impossible to disentangle the effects of UG and those of L1 (Hale, 1996; Yusa, 1999; Belikova & White, 2009), and it becomes a moot question to ask whether unconscious knowledge of structure dependence is derived from UG or L1. The principle of structure dependence plays a role in some constructions in one language but not in another. Question formation in English requires structure dependence because it moves elements (*wh*-phrases or auxiliaries), whereas the principle is inert in Japanese because the language forms questions by adding question particles rather than moving elements. This does not say that Japanese lacks structure dependence. The principle manifests itself in a new domain in Japanese grammar (i.e., the interpretation of multiple complex *wh*-questions; Nishigauchi, 1990). It might be claimed that our participants tapped knowledge of structure dependence via other structures in Japanese and transferred Japanese-based knowledge of the principle to totally new constructions, including the NI construction that does not occur in L1 (White, 2003). It is, however, unclear whether we receive enough evidence in our mother tongue to infer abstract properties of structure dependence, such as the dependence of linguistic rules on hierarchical syntactic structures. Moreover, the sample data for rejecting a structure-independent (linear) rule (i.e., negative evidence) is not accessible in L1 data, so that it is quite possible for a person to go through life without encountering any of the relevant examples that would choose between the structure-dependent rules and structure-independent rules (Berwick & Chomsky, 2008). Therefore, it is quite unlikely that the participants in our study tapped the properties of structure dependence via constructions in Japanese. Whether knowledge of structure dependence is due to UG or to L1, the present article shows that the principle of structure dependence, one of the core principles of UG, still functions in L2 acquisition and makes it possible for L2 learners to know more than is taught, which strongly argues against the critical period hypothesis.

The neuroimaging evidence that the principle of structure dependence is still available in second-language acquisition as well as L1 acquisition seems to be inconsistent with the selective recovery of one language in bilingual aphasia (see Paradis, 2004 for a variety of recovery patterns in aphasia). The selective recovery of one language in bilingual aphasia is often cited as evidence for the localization view that L1 and L2 are processed by different neural mechanisms (Albert & Obler, 1978). Generalizations from lesion research to healthy individuals are limited because of the compounding variables inherent in the lesion data (Perani & Abutalebi, 2005; Green & Price, 2001). The finding in this article is, however, consistent with a "dynamic view" of selective language recovery in bilingual aphasics: The selective language recovery does not reflect a different grammatical system between L1 and L2 but an impairment to the circuits involved in language control, that is, the ability of the system to choose one language over another (Abutalebi, Rosa, Tettamanti, Green, & Cappa, 2009; Green & Abutalebi, 2008).

Recent studies on L2 near-native speakers show that "failure" in L2 acquisition or divergence from native speakers is selective: Syntax is the locus of success; morphology or interface between syntax and phonology is the locus of failure (Sorace, 2005; White, 2003). In other words, purely syntactic properties can be completely acquired, whereas properties demanding the interaction of syntax and other domains (phonology, discourse/pragmatics) are resistant to complete mastery or native performance even after long-term exposure to them. With respect to those domains of grammar that are permanently defective in L2 acquisition, it is noteworthy that a long period of residence in a non-L1-speaking environment affects the L1 knowledge of the same domains in grammar (L2 influence on L1 attrition) (Sorace, 2005). In the broadest terms, general failure in adult L2 acquisition is only limited to the area requiring the integrated knowledge of syntax and discourse/pragmatics (e.g., articles) and that of syntax and morphology/phonology (e.g., subject–verb agreement). In contrast to the divergence from native speakers, no difference between native and L2 speakers can be found in the purely syntactic areas (i.e., the formation of the canonical word order and wh-interrogative sentences). Structure dependence belongs to the core computational aspect of syntax so that structure dependence may be immune to dysfunction in L2 acquisition even after puberty.

Our result shows that F3t serves the acquisition of a new syntactic rule, but this does not necessarily mean that Broca's area is exclusively dedicated to syntactic computation, considering the fact that Broca's area is reported to be active in nonsyntactic tasks in natural language (Grewe, Bornkessel, Zysset, & Wiese, 2006; Grodzinsky & Amunts, 2006; Grodzinsky & Friederici, 2006; Embick & Poeppel, 2005; Gelfand & Bookheimer, 2003) and syntactic tasks in nonnatural language (Friederici et al., 2006; Petersson, Forkstam, & Ingvar, 2004; Opitz & Friederici, 2003), and that structure dependence is omnipresent in music (Patel, 2003), in mathematical calculations (Varley, Klessinger, Romanowski, & Siegal, 2005), and in other cognitive domains (Tettamanti et al., 2009). Structure dependence might be domain-general and used for other functions, such as building hierarchically defined linguistic and nonlinguistic relations (Chomsky, 2004a; Marcus et al., 2003; Hauser et al., 2002). Further studies on structure dependence in other domains such as phonology, music, and mathematics are required to clarify the nature of structure dependence and language design in general, and the extent to which structure dependence is specialized for language. Structure dependence and the hierarchical structured representations (i.e., domain-specificity for language) might be derived from "some special arrangement of elements" in domain-general mechanisms and/or the way they interact with other cognitive systems (Chomsky, 2004b), but even assuming all these, L2 learners once again "face the same competing choice" between the structure-dependent rule and the structure-independent rule (for recent flawed challenges against the PoS argument, see Berwick & Chomsky, 2008). Apparently, some aspects of brain structure "predispose" L2 learners to follow the principle of structure dependence. Whatever status the principle of structure dependence is (for the specificity of structure dependence and recursive grammar in human language, see Piattelli-Palmarini et al., 2009), it is nonetheless worth noting that adult L2 learners in EFL contexts can use the principle and acquire more than is taught, contrary to what critical period accounts expect.

In summary, this study presents the very first neuroimaging evidence that L2 learners can acquire more syntactic knowledge than has been instructed, which suggests the possibility that L2 acquisition is guided by the interplay of nature (genetics, endowment, UG) and nurture (environment, instruction, learning), not by the unfortunate classic dichotomy "nature versus nurture." It also shows that at least the core principle (i.e., structure dependence) in the computational system of syntax proper still functions after the so-called critical or sensitive period, suggesting plasticity for the neural circuits for language in adult L2 learners. Finally, the present study indicates that classroom-based L2 instruction in EFL settings can cause changes in the brains of postpuberty learners (Sakai et al., 2009; Osterhout et al., 2008; Tatsuno & Sakai, 2005; McLaughlin, Osterhout, & Kim, 2004), supporting the view that the brain remains plastic or "trainable" through life, at least in some domains (Blakemore & Frith, 2005).

## Acknowledgments

We thank Hideaki Koizumi for his encouragement, Motoaki Sugiura for his technical advice, and Sanae Yamaguchi for her technical assistance. This work has been supported by a Grant-in-Aid for the promotion of "Brain Science and Education, Type II" from the Research Institute of Science and Technology for Society, Japan Science and Technology Agency (RISTEX/JST) to H. H., a Grant-in-Aid (S, 22222001) for Scientific Research from the Ministry of Education, Culture, Sports, Science, and Technology (MEXT) of Japan to M. K., and by Grants-in-Aid for Scientific Research (C, 17520274) from the Japan Society for the Promotion of Science and Scientific Research on Priority Areas ("Higher-Order Brain Functions," 20020022) from the MEXT of Japan to N. Y.

Reprint requests should be sent to Professor Noriaki Yusa, Department of English, Miyagi Gakuin Women's University, 9-1-1 Aoba-ku, Sendai-shi, Miyagi 981-8557, Japan, or via e-mail: n_yusa@me.com.

## REFERENCES

Abutalebi, J., Rosa, P. A., Tettamanti, M., Green, D. W., & Cappa, S. F. (2009). Bilingual aphasia and language control: A follow-up fMRI and intrinsic connectivity study. *Brain and Language, 109,* 141–156.

Albert, M. L., & Obler, L. K. (1978). *The bilingual brain*. New York: Academic Press.

Bachrach, A. (2008). *Imaging neural correlated of syntactic complexity in a natural context*. MIT Ph.D. dissertation.

Bahlmann, J., Schubotz, R., & Friederici, A. D. (2008). Hierarchical artificial grammar processing engages Broca's area. *Neuroimage, 42,* 525–534.

Belikova, A., & White, L. (2009). Evidence for the fundamental difference hypothesis or not? Island constraints revisited. *Studies in Second Language Acquisition, 31,* 199–223.

Ben-Shachar, M., Palti, D., & Grodzinsky, Y. (2004). Neural correlates of syntactic movement: Converging evidence from two fMRI experiments. *Neuroimage, 21,* 1320–1336.

Berwick, R., & Chomsky, N. (2008). "Poverty of the stimulus" revisited: Recent challenges reconsidered. *Proceedings* *of the 30th Annual Conference of the Cognitive Science Society,* 383.

Birdsong, D. (2006). Age and second language acquisition and processing: A selective overview. *Language Learning, 56,* 9–49.

Blakemore, S.-J., & Frith, U. (2005). *The learning brain: Lessons for education*. Oxford: Blackwell.

Bley-Vroman, R. (1990). The logical problem of foreign language learning. *Linguistic Analysis, 20,* 3–49.

Caplan, D., Waters, G., & Alpert, N. (2003). Effects of age and speed of processing on rCBF correlates of syntactic processing in sentence comprehension. *Human Brain Mapping, 19,* 112–131.

Carruthers, P., Laurence, S., & Stitch, S. (2005). *The innate mind*. Oxford: Oxford University Press.

Cavanna, A. E., & Trimble, M. R. (2006). The precuneus: A review of its functional anatomy and behavioural correlates. *Brain, 129,* 564–583.

Chomsky, N. (1957). *Syntactic structures*. The Hague: Mouton.

Chomsky, N. (1980). *Rules and representations*. New York: Columbia University Press.

Chomsky, N. (2000). *New horizons in the study of language*. Cambridge, MA: Cambridge University Press.

Chomsky, N. (2004a). *The generative enterprise revisited*. Berlin: Mouton de Gruyter.

Chomsky, N. (2004b). Beyond explanatory adequacy. In A. Belletti (Ed.), *Structures and beyond* (pp. 104–131). Oxford: Oxford University Press.

Crain, S., & Nakayama, M. (1987). Structure dependence in grammar formation. *Language, 63,* 522–543.

Crain, S., & Thornton, R. (1998). *Investigations in Universal Grammar: A guide to experiments on the acquisition of syntax*. Cambridge, MA: MIT Press.

DeKeyser, R. M. (2000). The robustness of critical period effects in second language acquisition. *Studies in Second Language Acquisition, 22,* 499–533.

Elman, J. L., Bates, E. A., Johnson, M., Karmiloff-Smith, A., Parisi, D., & Punkett, K. (1996). *Rethinking Innateness: A connectionist perspective on development*. Cambridge, MA: MIT Press.

Embick, D., & Poeppel, D. (2005). Mapping syntax using imaging: Prospects and problems for the study of neurolinguistic computation. In K. Brown (Ed.), *Encyclopedia of language and linguistics* (2nd ed., pp. 484–486). Oxford: Elsevier.

Fitch, W. T. (2010). Three meanings of "recursion": Key distinctions for biolinguistics. In R. K. Larson, V. Déprez, & H. Yamakido (Eds.), *The evolution of human language* (pp. 73–90). Cambridge, UK: Cambridge University Press.

Friederici, A. D., Bahlmann, J., Heim, S., Schubotz, R., & Anwander, A. (2006). The brain differentiates human and non-human grammars: Functional localization and structural connectivity. *Proceedings of the National Academy of Sciences, U.S.A., 103,* 2458–2463.

Gelfand, J., & Bookheimer, S. (2003). Dissociating neural mechanisms of temporal sequencing and processing phonemes. *Neuron, 38,* 831–842.

Gentner, T. Q., Fenn, K. M., Margoliash, D., & Nusbaum, H. C. (2006). Recursive syntactic pattern learning by songbirds. *Nature, 440,* 1204–1207.

Green, D. W., & Abutalebi, J. (2008). Understanding the link between bilingual aphasia and language control. *Journal of Neurolinguistics, 21,* 558–576.

Green, D. W., & Price, C. (2001). Functional imaging in the study of recovery patterns in bilingual aphasics. *Bilingualism: Language and Cognition, 4,* 191–201.

Grewe, T., Bornkessel, I., Zysset, S., & Wiese, R. (2006). Linguistic prominence and Broca's area: The influence of animacy as a linearization principle. *Neuroimage, 32,* 1395–1402.

Grodzinsky, Y., & Amunts, K. (2006). *Broca's region*. Oxford: Oxford University Press.

Grodzinsky, Y., & Friederici, A. D. (2006). Neuroimaging of syntax and syntactic processing. *Current Opinion in Neurobiology, 16,* 1–7.

Hale, K. (1996). Can UG and L1 be distinguished in L2 acquisition? *Behavioral and Brain Sciences, 19,* 728–730.

Hauser, M. D. (2010). On obfuscation, obscurantism, and opacity: Evolving conceptions of the faculty of language. In R. K. Larson, V. Déprez, & H. Yamakido (Eds.), *The evolution of human language* (pp. 91–99). Cambridge, UK: Cambridge University Press.

Hauser, M. D., Chomsky, N., & Fitch, W. T. (2002). The faculty of language: What is it, who had it, and how did it evolve? *Science, 298,* 1569–1579.

Hawkins, R. (2000). Persistent selective fossilization in second language acquisition and the optimal design of the language faculty. *Essex Research Reports in Linguistics, 34,* 75–90.

Indefrey, P. (2006). A meta-analysis of hemodynamic studies on first and second language processing: Which suggested differences can we trust and what do they mean? In M. Gulburg & P. Indefrey (Eds.), *The cognitive neuroscience of second language acquisition* (pp. 279–304). Oxford, UK: Blackwell.

Johnson, J. S., & Newport, E. L. (1989). Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology, 21,* 60–90.

Kotz, S. A. (2009). A critical review of ERP and fMRI evidence on L2 syntactic processing. *Brain and Language, 109,* 68–74.

Makuuchi, M., Bahlmann, J., Anwander, A., & Friederici, A. D. (2009). Segregating the core computational faculty of human language from working memory. *Proceedings of the National Academy of Sciences, U.S.A., 106,* 8362–8367.

Marcus, G. F., Vouloumanos, A., & Sag, I. A. (2003). Does Broca's play by the rules? *Nature Neuroscience, 6,* 651–657.

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

Moro, A. (2008). *The boundaries of Babel: The brain and the enigma of impossible languages*. Cambridge, MA: MIT Press.

Musso, M., Moro, A., Glauche, V., Rijntjes, M., Reichenbach, J., Büchel, C., et al. (2003). Broca's area and the language instinct. *Nature Neuroscience, 6,* 774–781.

Nishigauchi, T. (1990). *Quantification in the theory of grammar*. Dordrecht: Kluwer Academic Publishing.

Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia, 9,* 97–113.

Opitz, B., & Friederici, A. D. (2003). Interactions of the hippocampal system and the prefrontal cortex in learning language-like rules. *Neuroimage, 19,* 1730–1737.

Opitz, B., & Friederici, A. D. (2004). Brain correlates of language learning: The neuronal dissociation of rule-based versus similarity-based learning. *Journal of Neuroscience, 29,* 8436–8440.

Osterhout, L., Poliakov, A., Inoue, K., McLaughlin, J., Valentine, G., Pitkanen, I., et al. (2008). Second language learning and changes in the brain. *Journal of Neurolinguistics, 21,* 509–521.

Paradis, M. (2004). *A neurolinguistic theory of bilingualism*. Amsterdam: John Benjamins Publishing.

Patel, A. D. (2003). Language, music, syntax and the brain. *Nature Neuroscience, 6,* 674–681.

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

Petersson, K. M., Forkstam, C., & Ingvar, M. (2004). Artificial syntactic violations activate Broca's region. *Cognitive Science, 28,* 383–407.

Piattelli-Palmarini, M., Uriagereka, J., & Salaburu, P. (2009). *Of minds and language: A dialogue with Noam Chomsky in the Basque country*. Oxford, UK: Oxford University Press.

Pinker, S. (1994). *Language Instinct: How the mind creates language*. New York: Harper Collins.

Reali, F., & Christiansen, M. (2005). Uncovering the statistical richness of the stimulus: A powerful cue for acquiring syntactic categories. *Cognitive Science, 22,* 425–469.

Sakai, K. L., Miura, K., Narafu, N., & Muraishi, M. (2004). Correlated functional changes of the prefrontal cortex in twins induced by classroom education of second language. *Cerebral Cortex, 14,* 1233–1239.

Sakai, L. K., Nauchi, A., Tatsuno, Y., Hirano, K., Muraishi, Y., Kimura, M., et al. (2009). Distinct roles of inferior frontal regions that explain individual differences in second language acquisition. *Human Brain Mapping, 30,* 2440–2452.

Samuels, R. (2004). Innateness in cognitive science. *Trends in Cognitive Sciences, 8,* 136–141.

Schwartz, B. D. (1998). Second language instinct. *Lingua, 106,* 133–160.

Shetreet, E., Friedmann, N., & Hadar, U. (2009). An fMRI study of syntactic layers: Sentential and lexical aspects of embedding. *Human Brain Mapping, 48,* 707–716.

Slobin, N. (2003). Negative inversion as nonmovement. *Syntax, 6,* 183–212.

Smith, N., & Tsimpli, I. A. (1996). *The mind of a savant*. Oxford, UK: Blackwell.

Sorace, A. (2005). Selective optionality in language development. In L. Cornips & K. P. Corrigan (Eds.), *Syntax and variation: Reconciling the biological and social* (pp. 55–80). Amsterdam: John Benjamins Publishing.

Talairach, J., & Tournoux, P. (1988). *Co-planar stereotaxic atlas of the human brain. 3-Dimensional proportional system: An approach to cerebral imaging*. New York: Thieme.

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

Tettamanti, M., Alkadhi, H., Moro, A., & Perani, D. (2002). Neural correlates for the acquisition of natural language syntax. *Neuroimage, 17,* 700–709.

Tettamanti, M., Rotondi, I., Perani, D., Scotti, G., Fazzio, F., Cappa, S. F., et al. (2009). Syntax without language: Neurobiological evidence for cross-domain syntactic computations. *Cortex, 45,* 825–838.

Tomasello, M. (2003). *Constructing a language: A usage-based theory of language*. Cambridge, MA: Harvard University Press.

Varley, R., Klessinger, N. J. C., Romanowski, C. A. J., & Siegal, M. (2005). Agrammatic but numerate. *Proceedings of the National Academy of Sciences, U.S.A., 102,* 3519–3524.

Vigneau, M., Beaucousin, V., Hervé, P. Y., Duffau, H., Crivello, F., Houdé, O., et al. (2006). Meta-analyzing left hemisphere language areas: Phonology, semantics, and sentence processing. *Neuroimage, 30,* 1414–1432.

Wartenburger, I., Heekeren, H. R., Abutalebi, J., Cappa, S. F., Villringer, A., & Perani, D. (2003). Early setting of grammatical processing in the bilingual brain. *Neuron, 37,* 159–170.

White, L. (2003). *Second language acquisition and universal grammar*. Cambridge, UK: Cambridge University Press.

Yusa, N. (1999). Multiple-specifiers and wh-island effects in second language acquisition. In E. Klein & G. Martohardjino (Eds.), *The development of second language grammars* (pp. 215–238). Mahwah, NJ: Lawrence Erlbaum.