

Activation changes of the left inferior frontal gyrus for the factors of construction and scrambling in a sentence

By Kyohei TANAKA,^{*1} Shinri OHTA,^{*1,†} Ryuta KINNO^{*2} and Kuniyoshi L. SAKAI^{*1,‡}

(Communicated by Masao ITO, M.J.A.)

Abstract: Our previous functional magnetic resonance imaging (fMRI) studies have indicated that the left dorsal inferior frontal gyrus (L. dF3op/F3t) and left lateral premotor cortex (L. LPMC) are crucial regions for syntactic processing among the syntax-related networks. In the present study, we further examined how activations in these regions were modified by the factors of construction and scrambling (object-initial type). Using various sentence types, we clarified three major points. First, we found that the main effects of construction and scrambling consistently activated the L. dF3op/F3t and L. LPMC. Secondly, the main effects of scrambling clearly localized activation in the L. dF3op/F3t and L. LPMC, indicating the more narrowed down processing of syntax. Thirdly, step-wise percent signal changes were observed in the L. dF3op/F3t, demonstrating synergistic effects of construction and scrambling. These results demonstrate the abstract and intensive nature of syntactic processing carried out by these regions, *i.e.*, the grammar center.

Keywords: language, fMRI, inferior frontal gyrus, lateral premotor cortex, construction, scrambling

Introduction

Natural languages, or human languages, have various universal properties, such as the presence of subject, object, and predicate in a sentence. The grammatical relation of each noun phrase (NP)

(“subject, direct object, or indirect object” in linguistic terms) is structurally determined,¹⁾ whereas, in Japanese, a case marker (nominative *-ga*, dative *-ni*, or accusative *-o*) of each NP is finalized at the surface level. On the other hand, a predicate independently assigns a semantic role (“agent, experiencer, or patient” in linguistic terms; an agent initiates the action, and an experiencer/patient is affected by the agent) to each NP. Moreover, these linguistic relationships for individual NPs are not the simple one-to-one correspondence. Based on Kuroda’s analyses of Japanese sentences [see Chapter 6 in Kuroda (1992)²⁾], *subject-initial* or normal sentence patterns with NPs are as follows:

- I Transitive sentence pattern: NP-*ga* NP-*o*
- Ia Transitive sentence pattern: NP-*ga* NP-*ni*
- II Ergative sentence pattern: NP-*ni* NP-*ga*
- III Intransitive sentence pattern: NP-*ga*

The proposal of the ergative pattern II conforms to the “Dative Subject Constructions”.³⁾ Given these theoretical studies, the neuroscience of language should elucidate the neural basis of linguistic structures/constructions as precise as possible. Functional magnetic resonance imaging (fMRI) has been widely used as the most powerful method of visualizing localized activations as a statistical map in the

^{*1} Department of Basic Science, Graduate School of Arts and Sciences, The University of Tokyo, Tokyo, Japan.

^{*2} Division of Neurology, Department of Internal Medicine, Showa University School of Medicine, Tokyo, Japan.

[†] Correspondence should be addressed to: K. L. Sakai, Department of Basic Science, Graduate School of Arts and Sciences, The University of Tokyo, 3-8-1 Komaba, Meguro-ku, Tokyo 153-8902, Japan (e-mail: sakai@mind.c.u-tokyo.ac.jp).

[‡] Present address: Department of Linguistics, Faculty of Humanities, Kyushu University, Fukuoka, Japan.

Abbreviations: Act/Pas/Pot: active/passive/potential conditions; Act+/Pas+/Pot+: scrambled active/passive/potential conditions; [–scrambling]: Act, Pas, and Pot conditions; [+scrambling]: Act+, Pas+, and Pot+ conditions; Easier cond.: Act, Act+, and Pas conditions; Harder cond.: Pas+, Pot, and Pot+ conditions; SC: sentence control; Cont: control; NP: noun phrase; RTs: reaction times; IES: inverse efficiency scores (RTs/Accuracy); FWE: family-wise error; BA: Brodmann’s area; L: left; M: medial; LPMC: lateral premotor cortex; F3op/F3t/F3O: opercular/triangular/orbital parts of the inferior frontal gyrus; dF3op/F3t: dorsal F3op/F3t; pre-SMA: pre-supplementary motor area; AG: angular gyrus; IPS/SMG: intraparietal sulcus/supramarginal gyrus; MTG: middle temporal gyrus; pMTG/ITG: posterior middle/inferior temporal gyrus; LG/FG: lingual/fusiform gyrus.



Fig. 1. A picture-sentence matching task and a Control task. Under the two-argument (*i.e.*, Act, Act+, Pas, Pas+, Pot, and Pot+) and one-argument (*i.e.*, SC) conditions, each stimulus consisted of one picture (*top*) and one sentence (*bottom*). Pictures consisted of two stick figures, each of which was distinguished by a “head” symbol: a circle, square, or triangle. We used four kinds of grammatical particles, which represent the syntactic information in Japanese: *-ga*, a nominative case marker; *-ni*, a dative case marker; *-o*, an accusative case marker; and *-to*, a coordinator (“and”). Under both the [-scrambling] (marked in orange) and [+scrambling] (marked in red) conditions, we tested three constructions: active, passive, and potential sentences. Examples of matched sentences are shown in the first and second rows. For examples of mismatched sentences, see Fig. 1A in Kinno *et al.* (2014).⁶⁾ Under the SC condition, examples of matched (left panel) and mismatched (middle panel) sentences are shown; for the Control task, a matched example is shown (right panel).

human brain. Our previous studies have clarified the existence of the *grammar center*, which consists of two left frontal regions: the opercular/triangular parts of the left inferior frontal gyrus (L. F3op/F3t) and the left lateral premotor cortex (L. LPMC).⁴⁾ According to studies using a picture-sentence matching task with short Japanese sentences, specific *syntactic* loads due to more complex sentence structures significantly activated the grammar center.⁵⁾ Moreover, this task was successfully used for patients with a left frontal glioma, revealing three syntax-related networks including the grammar center.⁶⁾ In those studies, three types of sentences were tested:

active (Act) (*e.g.*, “○-*ga* □-*o* *oshiteru*”), passive (Pas) (*e.g.*, “□-*ga* ○-*ni* *osareru*”), and *scrambled* active (Act+) sentences (*e.g.*, “□-*o* ○-*ga* *oshiteru*”). In this article, we refer to “object scrambling” as simply scrambling, where an object to be emphasized is moved to the initial position of a sentence. In addition to these three sentence types, the present fMRI study tested scrambled passive (Pas+) (*e.g.*, “○-*ni* □-*ga* *osareru*”), potential sentences (Pot) (*e.g.*, “○-*ni* □-*ga* *oseteru*”), and scrambled potential sentences (Pot+) (*e.g.*, “□-*ga* ○-*ni* *oseteru*”) as well (Fig. 1, Table 1). By comparing these six conditions of constructions, we expected new *synergistic* effects

Table 1. Sentences used in this study

-Scrambling		+Scrambling	
Construction	English translation	Construction	English translation
Active sentence (Act)		Scrambled active sentence (Act+)	
○-ga □-o hiiteru	○ pulls □	□-o ○-ga hiiteru	As for □, ○ pulls it
○-ga □-o oshiteru	○ pushes □	□-o ○-ga oshiteru	As for □, ○ pushes it
○-ga □-o shikatteru	○ scolds □	□-o ○-ga shikatteru	As for □, ○ scolds it
○-ga □-o ketteru	○ kicks □	□-o ○-ga ketteru	As for □, ○ kicks it
○-ga □-o tataiteru	○ hits □	□-o ○-ga tataiteru	As for □, ○ hits it
○-ga □-o yonderu	○ calls □	□-o ○-ga yonderu	As for □, ○ calls it
Passive sentence (Pas)		Scrambled passive sentence (Pas+)	
□-ga ○-ni hikareru	□ is affected by ○'s pulling it	○-ni □-ga hikareru	As for ○'s pulling, □ is affected
□-ga ○-ni osareru	□ is affected by ○'s pushing it	○-ni □-ga osareru	As for ○'s pushing, □ is affected
□-ga ○-ni shikarareru	□ is affected by ○'s scolding it	○-ni □-ga shikarareru	As for ○'s scolding, □ is affected
□-ga ○-ni kerareru	□ is affected by ○'s kicking it	○-ni □-ga kerareru	As for ○'s kicking, □ is affected
□-ga ○-ni tatakareru	□ is affected by ○'s hitting it	○-ni □-ga tatakareru	As for ○'s hitting, □ is affected
□-ga ○-ni yobareru	□ is affected by ○'s calling it	○-ni □-ga yobareru	As for ○'s calling, □ is affected
Potential sentence (Pot)		Scrambled potential sentence (Pot+)	
○-ni □-ga hiketeru	○ can pull □	□-ga ○-ni hiketeru	As for □, ○ can pull it
○-ni □-ga oseteru	○ can push □	□-ga ○-ni oseteru	As for □, ○ can push it
○-ni □-ga shikareteru	○ can scold □	□-ga ○-ni shikareteru	As for □, ○ can scold it
○-ni □-ga kereteru	○ can kick □	□-ga ○-ni kereteru	As for □, ○ can kick it
○-ni □-ga tataketeru	○ can hit □	□-ga ○-ni tataketeru	As for □, ○ can hit it
○-ni □-ga yobeteru	○ can call □	□-ga ○-ni yobeteru	As for □, ○ can call it
Sentence control (SC)		English translation	
○-to □-ga neteiru		○ and □ lie	
○-to □-ga tatteru		○ and □ stand	
○-to □-ga aruiteru		○ and □ walk	
○-to □-ga hashitteru		○ and □ run	
○-to □-ga koronderu		○ and □ tumble	
○-to □-ga naiteru		○ and □ cry	

Only typical examples are shown here; other combinations of the symbols ○, □, and △ were also used. For brevity, English translation for the active and potential sentences was written in the present tense, although the tense of the original sentences was the progressive. -ga, nominative case marker; -o, accusative case marker; -ni, dative case marker; and -to, coordinator “and”.

on the cortical activation, which was not possible with our previous paradigm. We then examined how activations in the regions including the grammar center were modified by the factors of construction [active, passive, potential sentences] and scrambling [-scrambling, +scrambling]. Because these factors are related with shared aspects of syntactic structures in a sentence, we hypothesized that the main effects of construction and scrambling would cause overlapped enhancement of activations in the grammar center. We believe that the present study advances our knowledge about the cortical localization of syntax, in that the overlapped region reliably identifies the most critical grammar center in the brain.

Materials and methods

Participants. We recruited 22 native Japanese speakers, none of whom had a history of neurological disorders. To ensure the reliable and stable execution of the tasks despite fatigue and/or lack of concentration, we set an inclusion criterion for the participants, such that the error rates of every condition were less than 30%; six participants were dropped for this reason. The remaining 16 participants (12 males and four females; mean ± standard deviation [SD] age [yrs]: 27 ± 6.6) showed right-handedness (laterality quotients: 89 ± 15) according to the Edinburgh inventory.⁷⁾ Prior to participation in the study, written informed consent was obtained

from each participant after the nature and possible consequences of the study were explained. Approval for the experiments was obtained from the institutional review board of the University of Tokyo, Komaba Campus.

Stimuli. The four normal sentence patterns (I–III) in the Introduction correspond to active, passive, potential, and sentence control (SC) conditions, respectively (Fig. 1). For the patterns I, Ia, and II, *object-initial* or scrambled sentences are also grammatical in Japanese. The inclusion of patterns Ia and II in our experiments controls any effects concerning the order of grammatical particles *-ga* and *-ni* (see Table 2). On the other hand, passiveness is marked by the verb form *-(r)are-ru*, whereas potentialness is clearly marked by the verb form *-e-ru*. Although the verb form *-(r)are-ru*, *e.g.*, *kanji-rare-ru*, has multiple functions of marking the passive voice (*be felt*), potential (*can feel*), spontaneous (*be spontaneously felt*), and honoric form (*would feel*),⁸⁾ our passive sentences were least ambiguous (see Table 1). For the active construction, we used an abbreviated form (*-te-ru*) of progressive tense for the verb morphology (*-te-iru*); for the potential construction, we also used an abbreviated form (*-e-te-ru*) of progressive tense for the verb morphology (*-e-te-iru*). This was done to equate the number of letters under all tested conditions. Utilizing this experimental paradigm, we would be able to focus on the natural process of syntactic judgment, thereby excluding the surface-level effects of grammatical features.

Each visual stimulus consisted of a picture with head symbols (○, □, or △) at the top, and of an always grammatical sentence at the bottom (Fig. 1). For each stimulus, we chose two different head symbols. The sentences describing actions were written using a combination of the hiragana and kanji writing systems. We used four kinds of grammatical particles, which represent the syntactic information in Japanese: *-ga*, *-ni*, *-o*, and *-to*, a coordinator (“*and*”).

As listed in Table 1, two sets of Japanese verbs (six transitive verbs: *pull*, *push*, *scold*, *kick*, *hit*, and *call*; and six intransitive verbs: *lie*, *stand*, *walk*, *run*, *tumble*, and *cry*) were used, each of which, including the passive forms, had either four or five syllables. Note that the verb “*call*” is used only as a transitive verb in Japanese. There was no significant difference in frequency between the two sets of verbs ($t(10) = 0.7$, $P = 0.5$), according to the Japanese lexical database (“*Nihongo-no Goitokusei*” (Lexical Properties of Japanese), Nippon Telegraph and Telephone Corporation Communication Science Laboratories, Tokyo, Japan, 2003). The numbers of syllables and letters were strictly controlled among all conditions.

Under the two-argument conditions (*i.e.*, Act, Act+, Pas, Pas+, Pot, and Pot+), in which each sentence ended with a transitive verb and had two arguments (phrases associated with the predicate), the number of lines used in each picture except the head symbols was 14 ± 2.4 . Under the one-argument (*i.e.*, SC) condition, equally complex pictures (number of lines, 14 ± 2.5) were used, in which each sentence ended with an intransitive verb and corresponded to a double subjects (double agents) type, which did not involve two-argument relationships. Under these conditions, half of the pictures depicted actions occurring from left to right, and the other half depicted actions occurring from right to left (see Fig. 1); head symbols were also counter-balanced for both sides. These pictures further excluded the involvement of pragmatic information about word use (*e.g.*, “*An officer chases a thief*” is more acceptable than “*A thief chases an officer*”). There were 24 different stimuli (*i.e.*, different combinations between pictures and sentences) under each of these conditions.

All stimuli were presented visually in yellow against a dark background (Fig. 1). Each stimulus was presented for 6-s (intratrial interval) followed by a 2-s blank interval. The intratrial interval was fixed for 6 s irrespective of participants’ responses. To minimize the effect of general memory demands, a whole sentence of a minimal length (*i.e.*, two noun phrases and a verb) was visually presented for an ample time for the participants to respond, even for patients who were to be tested in a subsequent study. The stimuli were thus more appropriate than sequentially presented stimuli that involve memorization. For fixation, a red cross was also shown at the center of the screen to initiate eye movements from the same fixed position, and the participants were instructed to return their eyes to this position after

Table 2. Linguistic relationships for individual nouns

Construction	Case markers	Grammatical relations	Semantic roles
Active (Act, Act+)	<i>-ga</i>	subject	agent
	<i>-o</i>	direct object	patient
Passive (Pas, Pas+)	<i>-ga</i>	subject	experiencer
	<i>-ni</i>	indirect object	agent
Potential (Pot, Pot+)	<i>-ni</i>	subject	agent
	<i>-ga</i>	direct object	patient

the response. The stimulus presentation and collection of behavioral data [error rate and reaction times (RTs)] were controlled using the LabVIEW software and interface (National Instruments, Austin, TX). The participants wore earplugs and an eyeglass-like MRI-compatible display (resolution, 800×600 , within the visual angle of 7°) (VisuaStim XGA; Resonance Technology Inc., Northridge, CA).

Tasks. In the picture-sentence matching task (Fig. 1), the participants read a sentence silently and judged whether or not the action depicted in a picture matched the meaning of the sentence. They were instructed to respond by pressing one of two buttons in a row (right for a matched pair, and left for a mismatched pair) with the right hand as quickly as possible while ensuring correct responses. Under the two-argument conditions, all mismatched sentences were made by exchanging two symbols in the original sentences, e.g., “ \square pushes \circ ” instead of “ \circ pushes \square ”. Under the SC condition, both symbol-mismatched sentences and action-mismatched ones were presented equally often, requiring the sentences to be read completely. The participants underwent short practice sessions before the task sessions to become fully familiarized with this task. For some participants, the stimuli under the potential conditions were relatively harder to comprehend. Before the experiment, we explained typical potential sentences such as “*Taroo-nii(-wa) eigo-ga hanas-e-ru*” (“*Taroo can speak English*”). We did not use the particle *-wa* for the subject, because topicalization with *-wa* had differential effects other than scrambling.⁹⁾

For the Control (Cont) task, using the same stimulus sets of pictures and letters presented under the conditions described above, the participants judged whether or not the two head symbols in the picture matched those at the bottom, irrespective of their order (Fig. 1). The letters in hiragana were jumbled without changing the head symbols and kanji, so that the letter string prevented even basic word recognition. General cognitive factors such as visual perception of the stimuli, matching, response selection, and motor response were controlled by using the SC condition. The Cont task was used to lower the baseline activity. A single run of the task sessions (256 s) contained 32 trial events (four for each of the Act, Act+, Pas, Pas+, Pot, Pot+, SC, and Cont task conditions), the order of which was pseudorandomized to prevent any condition-specific strategy. Half of the stimuli consisted of matched picture-sentence pairs, and the other half consisted of mismatched pairs. Six runs were performed and

the participants did not encounter the same sentence twice.

MRI data acquisition. For the MRI data acquisition, the participant was in a supine position, and his or her head was immobilized inside the radio-frequency coil. The MRI scans were conducted on a 3.0 T system (GE Signa HDxt 3.0 T; GE Healthcare, Milwaukee, WI). We scanned 30 horizontal slices, each 3-mm thick and having a 0.5-mm gap, covering the range of $z = -38.5$ to 66 mm from the anterior to posterior commissure (AC-PC) line in the vertical direction, using a gradient-echo echo-planar imaging (EPI) sequence (repetition time (TR) = 2 s, echo time (TE) = 30 ms, flip angle (FA) = 90° , field of view (FOV) = $192 \times 192 \text{ mm}^2$, resolution = $3 \times 3 \text{ mm}^2$). In a single run, we obtained 128 volumes following four dummy images, which allowed for the rise of the MR signals. After completion of the fMRI session, high-resolution T1-weighted images of the whole brain (192 axial slices, $1.0 \times 1.0 \times 1.0 \text{ mm}^3$) were acquired from all participants with a three-dimensional fast spoiled gradient recalled acquisition in the steady state (3D FSPGR) sequence (TR = 8.4 ms, TE = 2.6 ms, FA = 25° , FOV = $256 \times 256 \text{ mm}^2$). These structural images were used for normalizing fMRI data.

fMRI data analyses. The fMRI data were analyzed in a standard manner using SPM12 statistical parametric mapping software (Wellcome Trust Center for Neuroimaging, <http://www.fil.ion.ucl.ac.uk/spm/>)¹⁰⁾ implemented on MATLAB (Math Works, Natick, MA). The acquisition timing of each slice was corrected using the middle slice (the 15th slice chronologically) as a reference for the EPI data. We realigned the time-series data in multiple runs to the first volume in all runs, and further realigned the data to the mean volume of all runs. The realigned data were resliced using seventh-degree B-spline interpolation, so that each voxel of each functional image matched that of the first volume.

After alignment to the AC-PC line, each participant's T1-weighted structural image was coregistered to the mean functional image generated during realignment. T1-weighted images were bias-corrected with light regularization, and segmented to the gray matter, white matter, cerebrospinal fluid, bone, other soft tissues, and air by using default tissue probability maps and the Segment tool in the SPM12, which uses an affine regularization to warp images to the International Consortium for Brain Mapping East Asian brain template.¹¹⁾ Inter-subject registration was achieved with Diffeomorphic Anatomical

Registration using the Exponentiated Lie algebra (DARTEL) toolbox in the SPM12.¹²⁾ The coregistered structural images were spatially normalized to the standard brain space as defined by the Montreal Neurological Institute (MNI) using DARTEL's Normalize to MNI Space tool. All of the normalized structural images were visually inspected and compared with the standard brain for the absence of any further deformation. The realigned functional images were also spatially normalized to the MNI space by using DARTEL's Normalize to MNI Space tool, which converted voxel sizes to $3 \times 3 \times 3 \text{ mm}^3$ and smoothed the images with an isotropic Gaussian kernel of 9-mm full-width at half maximum. Low-frequency noise was removed by high-pass filtering at 1/128 Hz.

In a first-level analysis (*i.e.*, the fixed-effects analysis), each participant's hemodynamic responses induced by the two-argument and SC conditions, as well as the Cont task, in each session were modeled with a boxcar function with a duration of 6 s from the onset of each visual stimulus. Because the maximum RTs for correct responses of each participant ranged from 3.8 s to 6.8 s, and 13 participants out of 16 took more than 5 s under some trials, we chose the whole presentation period of 6 s for analyses. The boxcar function was then convolved with a hemodynamic response function. Only the functional data for trials with correct responses were modeled. To minimize the effects of head movement, the six realignment parameters obtained from preprocessing were included as a nuisance factor in a general linear model. The images under each of the two-argument conditions, minus those under the SC condition, were then generated in the general linear model for each participant and used for the intersubject, across-subject, comparison in a second-level analysis (*i.e.*, the random-effects analysis). To examine the activation of the regions in an unbiased manner, we adopted whole-brain analyses.

A repeated-measures analysis of variance (rANOVA) with *t*-tests was performed with two factors (construction \times scrambling), the results of which were thresholded at uncorrected $P < 0.001$ for the voxel level, and at corrected $P < 0.05$ for the cluster level, with family-wise error (FWE) correction across the whole brain. For the anatomical identification of activated regions, we basically used the Anatomical Automatic Labeling method (<http://www.gin.cnrs.fr/AAL2/>)¹³⁾ and the labeled data as provided by Neuromorphometrics Inc. (<http://Neuromorphometrics.com/>) under academic sub-

scription. For each region of interest, we extracted the mean percent signal changes for each participant from the local maxima of each region using the MarsBaR-toolbox (<http://marsbar.sourceforge.net/>).

Results

Behavioral data. The error rate and RTs are shown in Fig. 2. For the error rate, an rANOVA with two factors (construction [active, passive, potential] \times scrambling [–scrambling, +scrambling]) showed that both main effects of construction ($F(2, 30) = 9.1$, $P = 0.0008$) and scrambling ($F(1, 15) = 6.5$, $P = 0.02$) were significant without the interaction of construction by scrambling ($F(2, 30) = 0.47$, $P = 0.9$) (Fig. 2A). These main effects were due to the fact that scrambled sentences were more difficult than the others, and the tasks became more difficult under the active, passive and potential conditions in this order.

Regarding RTs, there were significant main effects of construction ($F(2, 30) = 52$, $P < 0.0001$) and scrambling ($F(1, 15) = 30$, $P < 0.0001$), as well as a significant interaction between them ($F(2, 30) = 7.3$, $P = 0.03$) (Fig. 2B). Based on these results, we designated Act, Act+, and Pas the easier conditions, and called Pas+, Pot, and Pot+ the harder conditions. Indeed, paired *t*-tests showed that the mean of the RTs of the harder conditions was significantly longer than that of the easier conditions ($t(47) = 12$, $P < 0.0001$). Potential effects of task difficulty on brain activations will be considered later in the Results section.

Syntax-related activations in localized regions. Noting the significant difference in the RTs between the harder and easier conditions, we first examined brain activations by the contrast [Harder cond.] – [Easier cond.] (Fig. 3A). The most prominent activation was observed in the left frontal cortex including the grammar center, *i.e.*, the L. F3op/F3t and L. LPMC. The overall activation was consistent with the results of the contrast [Two-argument – Control] reported previously;⁶⁾ additional activation was observed in the thalamus, caudate, and putamen, as well as in the midbrain (Table 3).

Next we examined the main effects of construction by the conjunction of contrasts [Passive – Active] and [Potential – Passive] (Fig. 3B). The most prominent activation was found in the L. F3op/F3t and L. LPMC, and ventrally extended to the L. F3O. The medial pre-supplementary motor area (pre-SMA) was also activated (Table 3). We further examined the main effects of scrambling, *i.e.*,

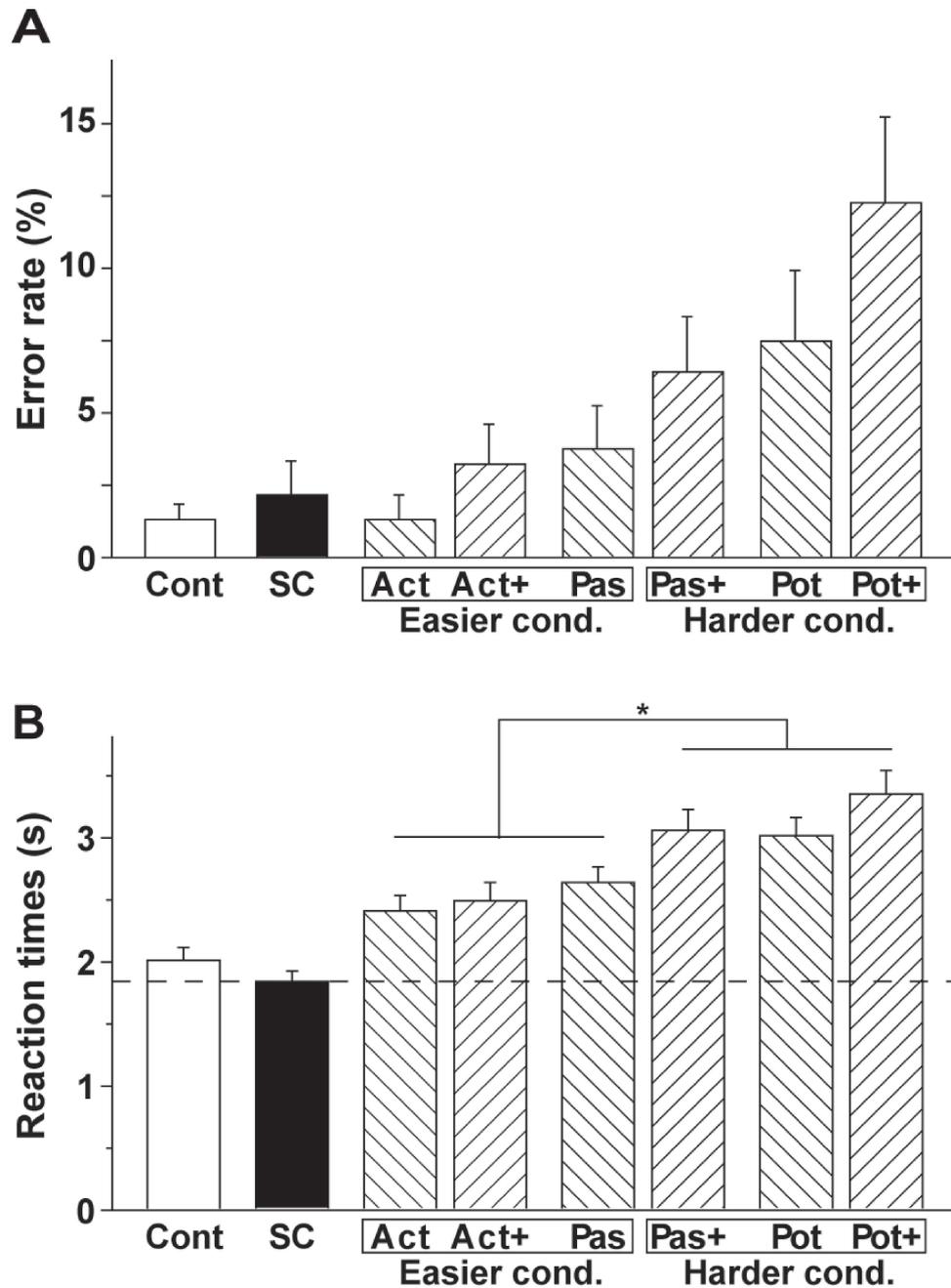


Fig. 2. Histograms for the error rate (A) and reaction times (RTs) (B). Error bars indicate the standard error of the mean for the participants. The RTs under the harder conditions (Pas+, Pot, and Pot+) were significantly longer than those under the easier conditions (Act, Act+, and Pas). An asterisk indicates $P < 0.0001$ (Paired t -test).

[+scrambling] – [–scrambling]. The activation was clearly localized in the L. dF3op/F3t and extended to the L. LPMC (Fig. 3C, Table 3), which was consistent with our hypothesis.

The cortical activations modified by syntactic loads. We further selected a local maximum

of the L. dF3op/F3t based on the focal activation shown in Fig. 3C and Table 3. Figure 4A shows the step-wise percent signal changes for the two-argument conditions, demonstrating synergistic effects of construction and scrambling, in the order of the active, passive and potential conditions, as well as in

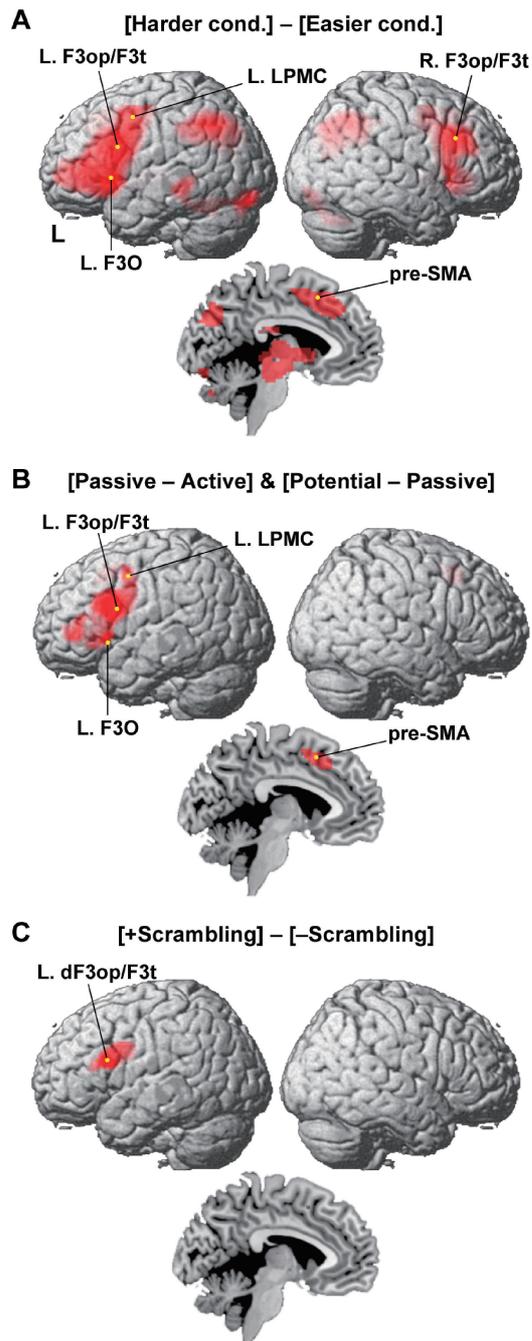


Fig. 3. Syntax-related activations in localized regions. (A) Cortical regions identified by [Harder cond.] – [Easier cond.]. Activations were projected onto the left (L) and right lateral surfaces of a standard brain (FWE-corrected $P < 0.05$). Medial sections are also shown. (B) Cortical regions identified by the conjunction of [Passive – Active] and [Potential – Passive]. Note the prominent activations in the left inferior frontal gyrus. (C) Cortical regions identified by [+scrambling] – [–scrambling]. Activations were further localized in the L. dF3op/F3t, extended to the L. LPMC.

the order of the [–scrambling] and [+scrambling]. We also selected each local maximum of the L. LPMC, L. F3O, and medial pre-SMA (left local maximum) based on the activation shown in Fig. 3B; the potential sentences induced significantly higher activations than the active sentences in these regions (Fig. 4B–D). We further selected a local maximum of R. F3op/F3t from the various activated regions shown in Fig. 3A, which showed significant activations only under the harder conditions (Fig. 4E). To examine any effects of task difficulty or general cognitive loads (other than linguistic effects) on the activations, we further tested whether there was a correlation between the performances and percent signal changes in the L. dF3op/F3t under the harder conditions, which was consistently activated by both main effects. To estimate the performances of individual participants, we used inverse efficiency scores ($\text{IES} = \text{RTs}/\text{Accuracy}$)¹⁴ under the harder conditions. As shown in Fig. 4F, no significant correlation was observed ($R = 0.32$, $P = 0.23$), excluding the possibility that activations were due to task difficulty. We conclude that the activations in the L. dF3op/F3t were mainly reflected the loads of syntactic processing.

Discussion

Using the various sentence types under the two-argument conditions, we clarified three major points. First, we found that the main effects of construction and scrambling, as well as the contrast [Harder cond. – Easier cond.], consistently activated the left frontal cortex including the L. dF3op/F3t and L. LPMC (Fig. 3). Secondly, the main effects of scrambling clearly localized activation in the L. dF3op/F3t and L. LPMC (Fig. 3C), indicating the more narrowed down processing of syntax. Thirdly, step-wise percent signal changes were observed in the L. dF3op/F3t (Fig. 4A), demonstrating synergistic effects of construction and scrambling, in the order of the active, passive, and potential conditions, as well as in the order of the [–scrambling] and [+scrambling]. Moreover, from those activations, we successfully excluded any effects of the order of grammatical particles, *-ga* and *-ni*, which would be finalized at the surface level. If the order of *-ga* and *-ni* were a key factor, then the responses to Pas and Pot+, or those to Pas+ and Pot, should have been comparable (see Table 1). These results demonstrate the abstract and intensive nature of syntactic processing carried out by the grammar center, which was much greater than expected from previous results.

Table 3. Cortical regions identified by the contrasts

Brain region	BA	Side	x	y	z	Z	Voxels
[Harder cond.] – [Easier cond.]							
LPMC	6/8	L	–42	2	47	6.7	4119
		L	–42	2	26	6.7	*
F3op/F3t	44/45	L	–48	11	23	7.0	*
F3t	45	L	–42	44	2	6.7	*
F3O	47	L	–33	26	–4	7.0	*
pre-SMA	6/8	M	–3	11	50	7.1	*
Thalamus		M	–9	–16	5	4.8	*
Caudate/Putamen		L	–15	8	5	4.9	*
Midbrain		M	–6	–19	–13	4.9	*
			6	–25	–13	4.6	*
IPS/AG	7/39	L	–33	–58	35	7.1	2007
		R	36	–58	41	6.0	*
IPS/AG/SMG	7/39/40	R	48	–34	44	3.3	*
Cuneus/Precuneus		M	–6	–76	35	4.9	*
LPMC	6/8	R	33	–1	47	5.2	1173
F3op/F3t	44/45	R	51	23	32	5.4	*
F3O	47	R	30	29	–1	6.8	*
MTG	21	L	–60	–37	–7	5.0	580
pMTG/ITG	37/19	L	–45	–49	–22	4.1	*
LG/FG	18/19	L	–24	–85	–19	5.5	*
LG/FG	18/19	R	24	–85	–13	4.1	272
Cerebellum		R	27	–64	–34	4.4	*
[Passive – Active] & [Potential – Passive]							
LPMC	6/8	L	–42	2	47	4.2	840
		L	–42	2	26	4.4	*
F3op/F3t	44/45	L	–45	11	23	4.7	*
		L	–54	26	14	3.4	*
		L	–45	44	2	4.2	*
		L	–51	17	–1	3.5	*
F3O	47	L	–33	26	–1	4.4	*
pre-SMA	6	M	–3	14	50	5.1	197
[+Scrambling] – [–Scrambling]							
LPMC	6/8	L	–42	–1	26	3.8	229
dF3op/F3t	44/45	L	–48	14	20	4.0	*
		L	–60	11	17	3.2	*

Stereotactic coordinates (x , y , z) in the MNI space are shown for each activation peak of Z values. The threshold is set at uncorrected $P < 0.001$ for the voxel level and at corrected $P < 0.05$ for the cluster level, with FWE correction across the whole brain. The region with an asterisk is included within the same cluster shown one row above.

Scrambling in Japanese does *not* change the grammatical relations and semantic roles of a sentence (see Table 2),^{15),16)} whereas there are information structure distinctions (*e.g.*, emphasis) related to scrambling. Regarding the signal enhancement for the L. dF3op/F3t under the [+scrambling] conditions, we propose that this can be estimated based on the syntactic structures of a sentence, more

specifically by the Degree of Merge (DoM), a key computational concept that we previously introduced.¹⁷⁾ The DoM is defined as the maximum depth of merged subtrees within a given tree structure of a sentence, thereby quantifying the complexity of that sentence. Indeed, the DoM was able to successfully account for syntax-selective activations in the L. dF3op/F3t and supramarginal gyrus (SMG).¹⁸⁾ As

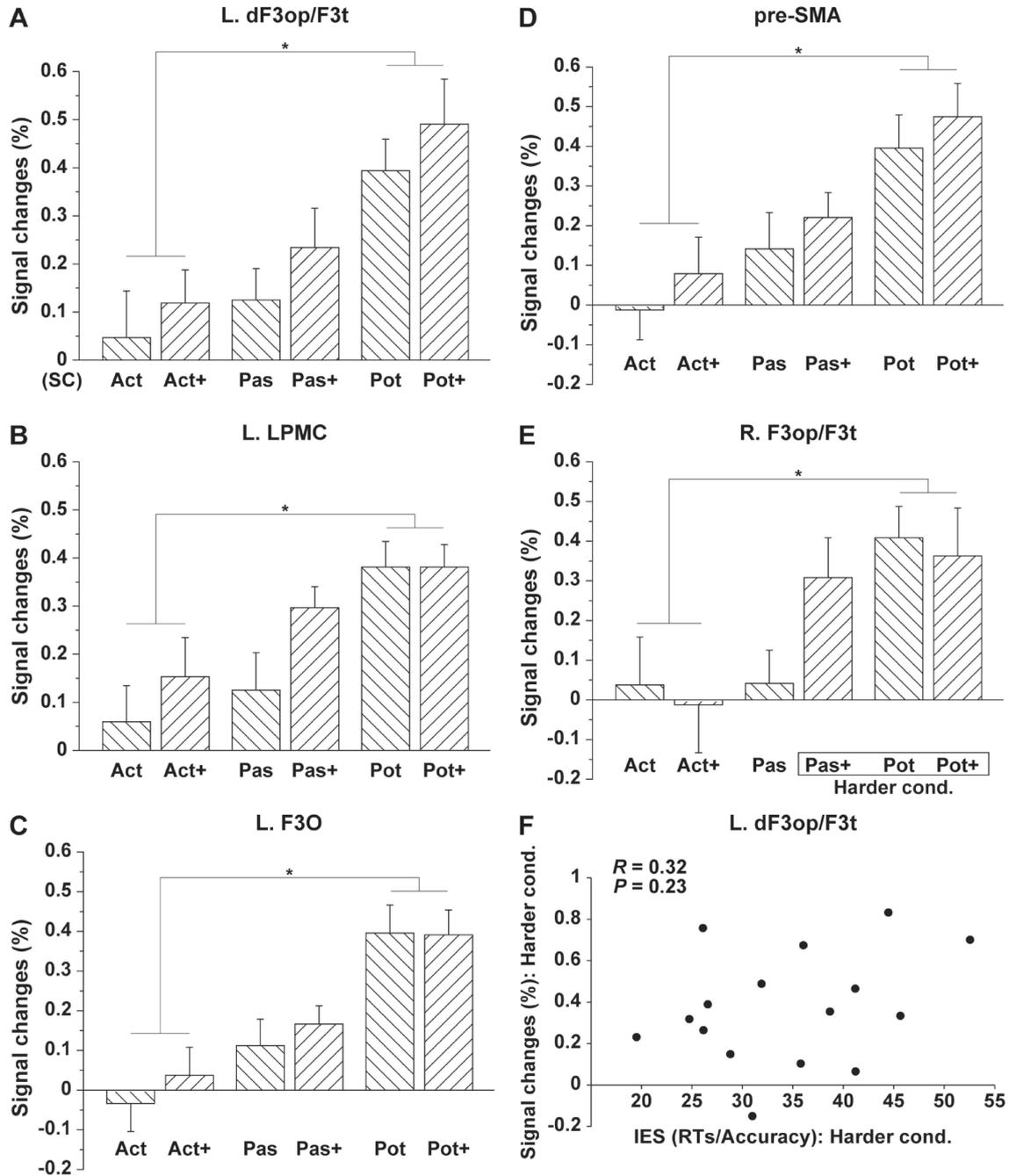


Fig. 4. The cortical activations modified by syntactic loads. Histograms for the percent signal changes are shown for each local maximum of the L. dF3op/F3t (A), L. LPMC (B), L. F3O (C), pre-SMA (D), and R. F3op/F3t (E). An asterisk indicates $P < 0.0001$ (Paired t -test). The signal changes are shown with reference to the SC condition. (F) The scatter plot for the percent signal changes in the L. dF3op/F3t and the IES (= RTs/accuracy) under the harder conditions, without a significant correlation.

discussed previously (see Fig. 6B in Ohta *et al.* (2013)¹⁷), the DoM of a scrambled sentence “ \bigcirc - o Δ - ga $hiiteru$ ” becomes three, while that of the corresponding active sentence “ Δ - ga \bigcirc - o $hiiteru$ ” becomes two. Likewise, the scrambled sentences

always yielded at least one more DoM than the corresponding passive or potential sentences, explaining the signal enhancement observed here.

It should be noted that the DoM of a passive sentence becomes larger than that of the corre-

sponding active sentence, because of forming a subordinate clause within the main clause “○-ga...-(r)are-ru”.¹⁷⁾ On the other hand, the DoMs under the Pas and Act+ conditions become comparable (see Figs. 6B and 7 in Ohta *et al.* (2013)¹⁷⁾), consistent with the similar signal changes observed under these conditions (Fig. 4A–D). As indicated by a linguistic study,²⁾ the base form of “*Taroo-ni(-wa) eigo-ga hanas-e-ru*” (“*Taroo can speak English*”) is an embedded sentence “*Taroo [Taroo eigo hanas] e-ru*”, explaining the larger DoM of the potential sentences than active sentences. In addition, potential sentences require an NP-deletion together with “Subject *ni*-Raising”,²⁾ which may explain the signal enhancement for the potential sentences relative to the passive sentences (Fig. 4A–D). The activated regions for the main effects of construction (Fig. 3B), which were much wider than those for the main effects of scrambling (Fig. 3C), indicate the presence of such additional syntactic processes. Moreover, the determination of the grammatical relations and semantic roles would become more intricate under the mixed conditions of the Act, Act+, Pas, Pas+, Pot, and Pot+ (see Table 2), which explains more enhanced activations than our previous work.⁵⁾

In a previous study using the picture-sentence matching task, we examined functional connectivity among 14 regions in the normal brain, and identified three syntax-related networks among these regions.⁶⁾ The five activated regions shown in Fig. 4 were all included in these networks. The L. dF3op/F3t, medial pre-SMA, and R. F3op/F3t were included in Network I, which is related to syntax and its supportive system; the L. LPMC was included in Network II, which plays a role of syntax and input/output interface for linguistic processing; and the L. F3O was included in Network III, which subserves the syntax-semantic interaction. Another patient study also indicated critical connections between the L. dF3op/F3t and the left intraparietal sulcus (L. IPS), between the L. LPMC and L. angular gyrus (AG), and between the L. F3t (ventral F3t) and L. F3O.¹⁹⁾ It is interesting to note that all of these regions were actually more activated under the newly introduced Harder conditions (Fig. 3A, Table 3). These results indicate that all of the three syntax-related networks contribute to the synergistic effects of construction and scrambling on the cortical activation.

We also found additional activations in the thalamus and caudate/putamen, as well as the midbrain by the contrast [Harder cond. – Easier

cond.]. The left caudate has been suggested to be sensitive to changes in the language or word meaning.²⁰⁾ In a study using depth recordings, the thalamus responded to syntactic or semantic errors in auditorily presented sentences.²¹⁾ Because the head of the caudate receives strong projections from the prefrontal cortex and thalamus, the cortico- and thalamic-caudate pathways may play a role in linguistic processing as well. In the present study, we clearly showed that activations of the syntax-related regions were modified by the two factors of construction and scrambling. It remains a question for future studies how the syntax-related networks are affected by lesions in the grammar center, when the same task and conditions are used. Such lesion studies could elucidate the causal link between activations and specific processes of syntax in the brain.

Acknowledgements

We thank N. Yusa for suggesting the use of potential sentences and other helpful input, N. Komoro for technical assistance, and H. Matsuda for administrative assistance.

References

- 1) Chomsky, N. (1981) Lectures on Government and Binding: The Pisa Lectures. Mouton de Gruyter, Berlin.
- 2) Kuroda, S.-Y. (1992) Japanese Syntax and Semantics: Collected Papers. Studies in Natural Language and Linguistic Theory. Kluwer Academic, Dordrecht.
- 3) Ura, H. (1999) Checking theory and dative subject constructions in Japanese and Korean. *J. East Asian Linguist.* **8**, 223–254.
- 4) Sakai, K.L. (2005) Language acquisition and brain development. *Science* **310**, 815–819.
- 5) Kinno, R., Kawamura, M., Shioda, S. and Sakai, K.L. (2008) Neural correlates of noncanonical syntactic processing revealed by a picture-sentence matching task. *Hum. Brain Mapp.* **29**, 1015–1027.
- 6) Kinno, R., Ohta, S., Muragaki, Y., Maruyama, T. and Sakai, K.L. (2014) Differential reorganization of three syntax-related networks induced by a left frontal glioma. *Brain* **137**, 1193–1212.
- 7) Oldfield, R.C. (1971) The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia* **9**, 97–113.
- 8) Shibatani, M. (1990) The Languages of Japan. Cambridge University Press, Cambridge, UK.
- 9) Ohta, S., Koizumi, M. and Sakai, K.L. (2017) Dissociating effects of scrambling and topicalization within the left frontal and temporal language areas: An fMRI study in Kaqchikel Maya. *Front. Psychol.* **8**, 748.

- 10) Friston, K.J., Holmes, A.P., Worsley, K.J., Poline, J.-P., Frith, C.D. and Frackowiak, R.S.J. (1995) Statistical parametric maps in functional imaging: A general linear approach. *Hum. Brain Mapp.* **2**, 189–210.
- 11) Ashburner, J. and Friston, K.J. (2005) Unified segmentation. *Neuroimage* **26**, 839–851.
- 12) Ashburner, J. (2007) A fast diffeomorphic image registration algorithm. *Neuroimage* **38**, 95–113.
- 13) Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., Mazoyer, B. and Joliot, M. (2002) Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage* **15**, 273–289.
- 14) Bruyer, R. and Brysbaert, M. (2011) Combining speed and accuracy in cognitive psychology: Is the inverse efficiency score (*IES*) a better dependent variable than the mean reaction time (*RT*) and the percentage of errors (*PE*)? *Psychol. Belg.* **51**, 5–13.
- 15) Fukui, N. (1993) Parameter and optionality. *Ling. Inq.* **24**, 399–420.
- 16) Saito, M. and Fukui, N. (1998) Order in phrase structure and movement. *Ling. Inq.* **29**, 439–474.
- 17) Ohta, S., Fukui, N. and Sakai, K.L. (2013) Computational principles of syntax in the regions specialized for language: Integrating theoretical linguistics and functional neuroimaging. *Front. Behav. Neurosci.* **7**, 204.
- 18) Ohta, S., Fukui, N. and Sakai, K.L. (2013) Syntactic computation in the human brain: The Degree of Merger as a key factor. *PLoS One* **8**, e56230.
- 19) Kinno, R., Ohta, S., Muragaki, Y., Maruyama, T. and Sakai, K.L. (2015) Left frontal glioma induces functional connectivity changes in syntax-related networks. *Springerplus* **4**, 317.
- 20) Crinion, J., Turner, R., Grogan, A., Hanakawa, T., Noppeney, U., Devlin, J.T., Aso, T., Urayama, S., Fukuyama, H., Stockton, K., Usui, K., Green, D.W. and Price, C.J. (2006) Language control in the bilingual brain. *Science* **312**, 1537–1540.
- 21) Wahl, M., Marzinzik, F., Friederici, A.D., Hahne, A., Kupsch, A., Schneider, G.-H., Saddy, D., Curio, G. and Klostermann, F. (2008) The human thalamus processes syntactic and semantic language violations. *Neuron* **59**, 695–707.

(Received Mar. 13, 2017; accepted May 26, 2017)