Check for updates

A Topographical Study on the Event-related Potential Correlates of Scrambled Word Order in Japanese Complex Sentences

Hiroko Hagiwara[1,2], Takahiro Soshi[1], Masami Ishihara[1], and Kuniyasu Imanaka[1]

Abstract

■ One of the most fundamental and universal properties of human language is a phenomenon called displacement. In the present study, we used multichannel event-related potentials (ERPs) to identify the nature of this phenomenon with Japanese, a subject-object-verb (SOV) language of relatively free word order. The ERPs of sentences of canonical word order (CC) were compared with those of non-canonical word order in two types of Japanese complex sentences; namely, in those which can be described as being in a middle-scrambled condition (MSC) and in those in a long-scrambled condition (LSC). The sustained anterior negativity (SAN) and the P600 in the pregap position were observed in the LSC, compared to the CC, and they are consistent with previous findings. The SAN, exhibiting a tripartite nature in morphology and scalp distribution, mainly reflected a storage cost of scrambled elements in sentence comprehension. The subsequent P600 had a left fronto-temporal maximum, distinguished from a posterior P600, taken as a reflector of the thematic role assignment in previous related studies. It is argued that the P600 in the present study reflects a cost of structural integration intensively depending on the case marker information. A compositional interpretation of sentence meanings was also observed, reflected in an anterior negativity at the postgap verbal position, which cannot be differentiated at the pregap verbal position in the languages of subject-verb-object (SVO) word order. ■

INTRODUCTION

The issue of what mental and neurological mechanisms are involved in the processing of displaced constituents, a most fundamental and universal property of human language, has been the focus of much investigation in recent psycholinguistic and neurolinguistic research in the domain of language comprehension. The most transparent and familiar manifestation of the displacement phenomena is observed in languages of rigid word order. In English, for example, the canonical word order is subject-verb-object (SVO), where the direct object follows the verb. In the comprehension of non-canonical interrogative sentences such as ''Which surgeon do you trust?,'' the processor must establish the syntactic dependency relation between the fronted *wh*-phrase, *which surgeon*, commonly referred to as the ''filler,'' and the verb *trust*, to identify *which surgeon* is the object of the sentence, referred to as the ''gap,'' a vacant position following *trust*, which has the semantic role, theme, of the verb. Cumulative evidence from processing studies of filler–gap dependencies of this type (i.e., a *wh*-filler and its gap) shows that the temporary storage of the displaced constituent, its retrieval from working memory, and its integration with a structural tree for semantic interpretation are all assumed to spend a certain amount of processing resources (Gibson, 1998, 2000). Subsequent electrophysiological studies, using event-related potentials (ERPs), suggest that these processes constitute separate components in sentence processing: The storage and the retrieval of the filler from working memory is reflected in a left anterior negativity (LAN), and the integration process in a late positivity, or P600.

The present study investigates whether a different type of filler–gap dependency, namely, a noun phrase as a filler (NP-filler) and its gap, elicits comparable ERP effects by utilizing Japanese scrambled sentences with noncanonical word order. Furthermore, we try to explore the possibility of differentiating the related ERP components into subcomponents and their corresponding functions, which are claimed to be associated with filler–gap dependency, using a multichannel electroencephalogram (EEG) system (128 channels).

In the ERP literature, that structural configuration-induced working memory costs was first reported by Kluender and Kutas (1993). On examining various types of filler–gap dependencies in English, they found a

[1]Tokyo Metropolitan University, [2]RISTEX, Japan Science and Technology Agency

negative-going voltage deflection with a left frontal maximum between 300 and 500 msec post-word-onset in response to a subject pronoun at the beginning of a *wh*-question, for example, *What have YOU*..., compared to the same pronoun at the beginning of a yes/no question, for instance, *Have YOU*.... They interpreted this transient LAN to be a reflection of the storage of a *wh*-filler in working memory.

A more slow and sustained LAN associated with the onset of the filler–gap dependency was uncovered by the recording of multiword ERPs. King and Kutas (1995) tested object and subject relative clauses in English and found a slow frontal negative potential in the object relative clause compared with the subject relative clause. This effect started at the position immediately after the introduction of the filler, a *wh*-pronoun, and continued to the verb of the object relative clause. They argued that, while analyzing upcoming syntactic structure, holding a *wh*-filler in working memory until the gap position is indexed by sustained anterior negativity (SAN). Exploring the distance effect of the filler–gap dependency in relation to individual differences in working memory capacity, Fiebach, Schlesewsky, and Friederici (2002) investigated the storage and integration processes of working memory in German indirect *wh*-questions. They also found a SAN starting at the left anterior focus at the prepositional phrase immediately following the *wh*-pronoun in an object *wh*-question, relative to a subject *wh*-question where the distance between filler and gap is large, in a group of subjects with low working memory capacity. They argue that this type of SAN reflects the additional memory resources required in an object *wh*-question (see also Felser, Clahsen, & Münte, 2003, for similar observations in German *wh*-questions). Phillips, Kazanina, and Abada (2005) more closely examined the distance effect of the filler–gap dependency by scrutinizing the locus of increased cost, and found that the storage cost is not correlated with distance.

In addition to the transient LAN and/or the SAN associated with filling gaps, some studies have reported a LAN effect *after* the gap has been filled. Kluender and Kutas (1993) compared the response to the preposition immediately after the embedded direct object gap and that of three words after the gap, as in ''*Can't you tell what she intends to drum* ____ *INTO you BY*...'' with the same prepositions in yes/no questions, as in ''*Can't you tell if she intends to drum this stuff INTO you BY*....'' They observed a LAN at the prepositions of both types with the latency of 300–500 msec most prominent at the left anterior-temporal electrode in the embedded object gap condition. They interpreted the LAN at the preposition immediately after the gap as a reflection of the retrieval of the filler from working memory, and that seen three words after the gap as ''the overall difficulty of processing sentences containing unbounded dependencies (p. 206).'' King and Kutas (1995) also observed a LAN effect at the verb of the matrix clause, which immediately follows the gap in the object relative sentence. This effect was superimposed on the slow bilateral negative potential ranging across the subsequent phrase. They attributed this effect to the difficulty in the process of thematic role assignment.

In short, these results indicate that a LAN is not exclusively related to the storage and retrieval of working memory, and suggest that several distinctive functions are reflected in the seemingly single component. One of the major objectives of this present study is to explore these possibilities.

In terms of word order variations induced by *wh*-movement, there is another type of dislocation operation called NP-movement, one example of which is known as scrambling. Scrambling is a term for the phenomenon referred to as free word order. There are only a few languages in the world which exhibit such an effect. German and Japanese are instances of these types of languages. Rösler, Pechmann, Streb, Roder, and Henninghausen (1998) examined the processing of German ditransitive sentences in which the subject (S), the indirect object (IO), and the direct object (DO) were presented in various scrambled orders in simple sentences. The ERPs revealed negative potentials with a left-anterior focus on the determiner of a scrambled NP with accusative case in DO–S–IO sentences or with dative case in IO–S–DO sentences, compared to the determiner denoting nominative case in canonical sentences of the order S–IO–DO. Although the authors argued that this effect is related to working memory load, insofar as a scrambled NP that cannot be immediately assigned to its canonical position must be held in working memory, some researchers suggested the possibility of the observed negativity as a reflection of a syntactic effect. Friederici, Schlesewsky, and Fiebach (2003) pointed out that the negative potential was restricted only to the determiner and did not extend to the object NP, thereby it resembled a transient LAN, which is reported to be elicited by various types of morphosyntactic violations (Coulson, King, & Kutas, 1998; Gunter, Stowe, & Mulder, 1997; Penke et al., 1997; Münte, Heinze, & Mangun, 1993; Kutas & Hillyard, 1983). They reinterpreted Rösler et al.'s results arguing that the observed transient negativity signaled a mismatch between a predicted structural position and the element encountered there. This hypothesis is supported by the study of Matzke, Mai, Nager, Russeler, and Münte (2002) on German scrambling. They examined, among other things, simple sentences with nominative–accusative and accusative–nominative word order, and reported a transient LAN at the fronted object NP in noncanonical sentences compared with canonical sentences. Schlesewsky, Bornkessel, and Frisch (2003) also reported a similar effect in noncanonical ditransitive sentences.

There could be several reasons for the lack of a sustained negativity in the noncanonical word order instances of the simple transitive and ditransitive sentences in these studies (cf. Matzke et al., 2002). For example, given that syntactic complexity influences the appearance of the ERP component, one could argue that the sentences used in all of these studies were not complex enough to elicit a sustained slow-potential effect, and therefore, a scrambled word order did not place a burden on working memory. If the constructions had been more complex, the scrambled word order would have elicited a sustained effect (*Complexity Hypothesis*).

Another reason could be related to the type of displaced elements. It could be possible that a sustained negativity appears when the parser encounters a *wh*-phrase but not an NP as the filler. In other words, the question arises as to whether or not this effect is modulated by the type of dependency. Processing *wh*-phrases requires an operator-variable dependency, in addition to simply holding the filler. Furthermore, contrary to an ordinary NP, a *wh*-expression may well function as a trigger for the looking up of its upcoming trace, or gap position. These differences in property may contribute to the increase of memory load in the filler–gap dependency of a *wh*-type, thereby yielding a slow sustained negative potential. The filler–gap dependency of an NP-type, on the other hand, does not convey such functions, and therefore, might not elicit a sustained effect (*Displacement Type Hypothesis*).

To this end, Japanese is a perfectly suitable language to test these hypotheses because unlike German, where scrambling is restricted to being clause-internal, Japanese allows unbounded long-distance clause-external scrambling, that is, an embedded object NP can be fronted to a sentence-initial position, crossing the subject of the matrix clause. One previous study (Ueno & Kluender, 2003) examined ERP evidence of scrambled word order. In this study, the processing of simple sentences with canonical and scrambled word order was tested in the form of both *wh*-questions and yes/no questions with demonstrative objects. They found bilateral slow anterior negative potentials between the filler and the gap not only in scrambled *wh*-questions but also in yes/no questions. Of note here is that demonstratives resemble *wh*-expressions in that they obligatorily call for coreferential NPs. An ordinary NP such as *a lawyer* and *a dentist*, on the other hand, is itself referential in nature and does not corefer to any other expression. Therefore, it is certainly necessary to examine the processing of sentences using ordinary NPs in long-distance clause-external scrambling for the purpose of exploring the universal psychophysiological nature of displacement phenomena.

The Present Study

The goal of the present study is to investigate whether scrambled word order with the long-distance filler–gap dependency of the NP-type elicits a SAN and/or a transient LAN at the related position in question. In addition, we analyze the functional significance and neuronal mechanisms of the related ERP components on the basis of the information of topographical distribution with regard to the syntactic processes of displacement phenomena. We used three types of critical sentences, one canonical and two types of noncanonical sentences: All were composed of a matrix clause and a complement clause. Examples are shown in Table 1. The sentence in the canonical condition (CC) is a typical example of a Japanese sentence with canonical word order and a

Table 1. Experimental Conditions and Their Corresponding Structures

Adv/PP	*NP1*	*NP2*	*NP3*		*VP1*	*VP2*
Canonical Condition (CC)						
$[_{CP}$Kaiken-de	$[_{TopP}$ shacho-**wa**	$[_{CP}$ $[_{TP}$ hisho-**ga**	bengoshi-**o**		sagasiteiru]] to	itta.]]
at the meeting	the president (TOP)	the secretary (NOM)	the lawyer (ACC)		was looking for (COMP)	said.
"At the meeting, the president said that the secretary was looking for the lawyer."						
Middle-scrambled Condition (MSC)						
$[_{CP}$Kaiken-de	$[_{TopP}$ shacho-**wa**	$[_{CP}$ $[_{TP}$ bengoshi-**o**$_i$	$[_{TP}$ hisho-**ga**	t_i	sagasiteiru]] to	itta]]
at the meeting	the president (TOP)	the lawyer (ACC)	the secretary (NOM)		was looking for (COMP)	said.
Long-scrambled Condition (LSC)						
$[_{CP}$Kaiken-de	$[_{TopP}$ bengoshi-**o**$_i$	$[_{TopP}$ shacho-**wa**	$[_{CP}$ $[_{TP}$ hisho-**ga**	t_i	sagasiteiru]] to	itta]]
at the meeting	the lawyer (ACC)	the president (TOP)	the secretary (NOM)		was looking for (COMP)	said.

TOP = topic marker; NOM = nominative marker; ACC = accusative marker; CP = complementizer phrase; TopP = topic phrase; TP = tense phrase; t_i = a gap position for the noun phrase of the same index.

complement clause. On the other hand, the sentence in the middle-scrambled condition (MSC) consists of a sentence with noncanonical word order in that, the NP, *bengoshi-o* (''the lawyer''), is moved from the object position in the complement clause, crossing the subject NP, and adjoined to the embedded clause. In the long-scrambled condition (LSC), the object NP of the complement clause moves out of an embedded clause, crossing two NPs, and is adjoined to the matrix clause, the so-called long-scrambling construction.

Based on the hypotheses and assumptions outlined above, a number of predictions follow. First, if the *Complexity Hypothesis* holds true, a SAN will appear at the position of the first NP and span over the second and third NPs in the LSC. The sentences in the MSC, on the other hand, will not elicit a sustained effect because the number of phrasal nodes is smaller in the MSC, hence, is less complex than those in the LSC.

Secondly, assuming that *wh*-movement and NP-movement are distinct phenomena, the *Displacement Type Hypothesis* predicts that none of the ERP effects overlap between *wh*-movement and NP-movement with respect to anterior negativity. In the previous studies, a sustained negativity has been reported repeatedly in *wh*-movement, as mentioned above, whereas only a transient LAN was elicited in NP-movement (Schlesewsky et al., 2003; Rösler et al., 1998, cf. Matzke et al., 2002). Given these observations, this hypothesis predicts that a sustained negativity will not be observed in the sentences of the LSC compared to those of the CC. Instead, it is predicted that at the first NP, the scrambled object of the LSC elicits a transient LAN, compared to the topic NP in the matrix clause of the CC.[1]

Thirdly, in the previous studies, sentences with filler–gap dependencies frequently elicited a late positivity at around the position where the filler is integrated into its original position, the gap, in the structural tree (Ueno & Kluender, 2003; Fiebach et al., 2002; Kaan, Harris, Gibson, & Holcomb, 2000). These pre- or postgap positions could either be NPs or verb phrases (VPs). Some researchers interpret this effect as the computational cost to integrate NPs, which are held in working memory, into a syntactic structure and to assign thematic roles to these NPs, known as the *Integration Hypothesis* (Gibson, 1998). Given that in the previous studies the P600 was reported only in the case of *wh*-movement, *not* in typical NP-movement, we predict a late positivity in the MSC and the LSC only when the SAN precedes it, as was the case with *wh*-movement.

Concerning the topographical distribution on the LAN, there exist contrasting views on its syntactic specificity. Some studies reported only a sustained effect (Felser et al., 2003; Fiebach et al., 2002; King & Kutas, 1995; Kluender & Kutas, 1993), whereas others observed only a transient effect (Schlesewsky et al., 2003; Rösler et al., 1998), and their distributions are far from constant. The present study was intended to obtain a denser spatial sampling of the surface frontal negativity distribution than has previously been described. To date, the great majority of published studies in syntactic processing in relation to working memory have sampled scalp fields with a relatively low density of spatial sampling, usually around 20 electrodes, although a few studies have used as many as 51 (Bornkessel, Fiebach, & Friederici, 2004; Fiebach et al., 2002). We measured 128-channel ERPs elicited by visual presentation. If we succeed in differentiating a sustained effect from a transient response in latency, as well as in topographical distribution, it would suggest that these negativities are the reflection of distinct phenomena and are instantiated in different neural mechanisms (Rugg & Coles, 1995).

METHODS

Participants

Fourteen participants (8 women) between the ages of 18 and 24 (mean age 20.36 years, $SD = 1.65$) were recruited from the student population of the Tokyo Metropolitan University, and paid JPN¥5,000 for approximately 3.5 to 4 hours of participation. All were neurologically normal, healthy, right-handed, native speakers of Japanese, and had normal or corrected-to-normal vision. Hand preference was assessed in all participants via the Edinburgh Handedness Inventory (Oldfield, 1971). The participants gave informed consent in written form before the experiment. The Human Subjects Ethics Committee of Tokyo Metropolitan University approved of all the procedures.

Materials

Eighty sentences for each of the three sentence types, illustrated in Table 1, were constructed using the sentence frame *AdvP* or *PP/NP1/NP2/NP3/VP1/VP2*, whereby the NPs were composed of matrix subjects with the topic marker *–wa* (TOP), complement subjects with the nominative marker *–ga* (NOM), and complement objects with the accusative marker *–o* (ACC). The embedded verb VP1 was a transitive verb and always preceded the matrix verb VP2, which subcategorized for a complement clause. All of the NPs (NP1, NP2, and NP3) were animate and could all equally take the role of agent or patient of the transitive verb VP1 and the role of agent in the matrix verb VP2. All the sentences were grammatical, meaningful, and semantically neutral in that they did not include any nonreversible sentences such as ''the police pursued a criminal,'' whose meaning has biases toward the thematic roles of the two NPs. The embedded verbs included transitive verbs with Yamato (Japanese) origin and light verbs with Sino-Japanese origin. Complex predicates, causative predicates, and verbs with passive morphemes were excluded. The NPs used were restricted to common nouns such as "the professor," "the doctor," "the actress," and "the journalist." No proper names were used. The topic marker *–wa* was attached to the matrix subject NP, instead of the nominative marker *–ga*, to avoid any extra processing difficulties which may be caused by two consecutive NPs with the same nominative marker *–ga*. The adverbial phrase (AdvP) or a postpositional phrase (PP) was positioned sentence-initially, which strictly modified the matrix predicate to avoid contaminating the ERP responses of the critical word, the first NP, with any nonlinguistic factors. The AdvPs and PPs specified either time or location, and all of them were equally placed so as not to provide any information about the content or type of the following sentence.

In addition, 100 filler sentences of various sentence types with different degrees in comprehension difficulties were prepared. Sentences varied in length from 4 to 7 phrases. They included simple transitive sentences with canonical word order and some with noncanonical word order, negation sentences with negative polarity items, complex sentences with subordinate clauses, and garden-path sentences with center-embedded relative clauses. Twenty sentences for each of the five sentence types were constructed. A total of 340 sentences were divided into five blocks of 68 sentences each. Each block contained 48 sentences of the critical conditions (i.e., the CC, the MSC, and the LSC). The sequence of items and fillers was pseudorandomized. The five blocks were presented in a different order to the participants.

The comprehension questions for each sentence addressed the content of the proposition. The questions focused on who did what to whom (i.e., the agent or the patient of action). Moreover, they were constructed in such a way that the thematic roles referred to each of the three NPs and the action to the two VPs an equal number of times. There were an equal number of yes and no questions as well. Because the questions were unpredictable, the participant had to fully process the meaning of a sentence to respond correctly so that a particular processing strategy could not be adopted.

Procedure

Before the ERP experiment, a short practice session was conducted. The stimuli sentences for training were prepared separately, including 10 sentences for each of the three conditions and 30 filler sentences, resulting in a total of 60 sentences.

Participants were seated in a comfortable chair approximately 1 meter in front of a cathode-ray tube (CRT) monitor. Sentences were presented visually for one *bunsetsu*, or phrase, at a time (i.e., in a phrase-by-phrase fashion at the center of the screen in a horizontal angle of 3.0° 3.6° and a vertical angle of 1.5°). The maximum number of letters presented per phrase was less than seven. Each trial was structured as follows. A fixation point appeared on the screen for 500 msec. Each phrase was presented for 400 msec, followed by an interstimulus interval of 500 msec. After the terminal word of a sentence, the screen was blank for 500 msec, followed by a fixation point of 100 msec. The screen was then blank for 800 msec, followed by a comprehension question and answer, which was presented in one frame.

Each sentence was followed by a comprehension question. The participants' task was to judge whether or not the question and the answer were correct in relation to the critical item by pressing either yes or no buttons. The main purpose of the comprehension task was to control for the correct processing of the stimuli, and only the sentences the participants comprehended correctly were rendered to the subsequent ERP data analyses. The comprehension question and answer disappeared after a response had been registered for a timeout of 3 sec. The next trial began 2500 msec after the participant's subsequent button press. The participants were asked to blink during the intertrial interval. They were also allowed to take rests for 5 to 10 min, or more if necessary, between the blocks.

EEG Recording

A continuous EEG was recorded from 123 Ag/AgCl sintered electrodes mounted on an elastic cap (Neuroscan, Quickcap). Figure 1 shows the positioning of the

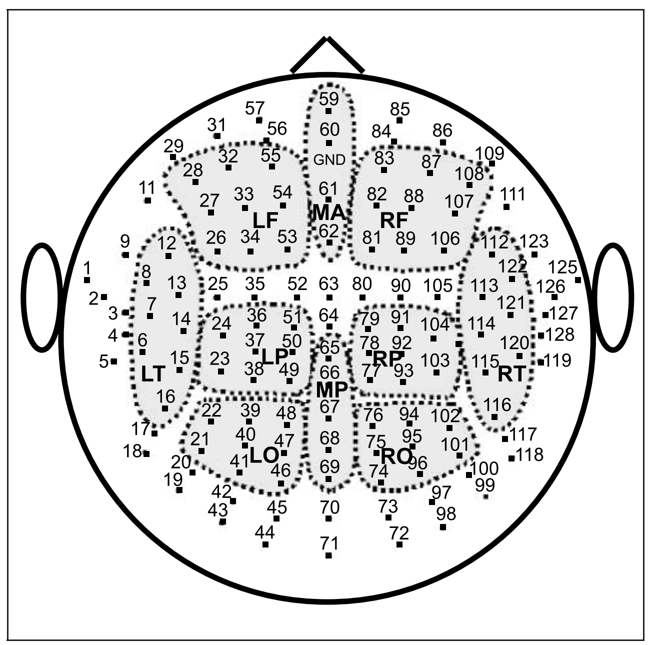


Figure 1. Diagram of 123 scalp electrode locations projected onto a two-dimensional surface used for portraying ERP topography. The electrode sites were divided into 10 regions, as separated by the dotted lines, which were submitted for statistical analysis. The abbreviations represent each region as follows: LF = left frontal; RF = right frontal; LT = left temporal; RT = right temporal; LP = left parietal; RP = right parietal; LO = left occipital; RO = right occipital; MA = mid anterior; MP = mid posterior.

electrodes. To reduce interparticipant variability of electrode positions, the Cz was positioned with anatomical measurements for each participant and, subsequently, electrode no. 63 was placed there. Other electrode locations were determined according to the geometry of the elastic cap on which the distance between electrodes is evenly distributed. This scheme is suitable for designating accurate topographical distribution of the ERP components. Other electrodes were posited at the outer canthus of each eye and beneath the left eye in order to derive off-line the vertical and horizontal electrooculograms (VEOG, HEOG) for artifact rejection of EEG epochs contaminated by eye movement and blink artifacts. All electrodes were referenced to the linked mastoids. Electrode impedance was kept below 5 kΩ through all recordings. The EEG was amplified with a bandpass from DC to 100 Hz and sampled at a frequency of 250 Hz.

Data Analysis

Behavioral Data

Comprehension accuracy and reaction times in the comprehension task were analyzed with a repeated measures analysis of variance (ANOVA), with Sentence Type representing a within-participant factor.

EEG Analysis

After the completion of the data collection, EEG files were segmented with respect to event markers into 2900 msec epochs (−200 to 2700 msec) for the multiphrase ERPs and 1100 msec epochs (−200 to 900 msec) for the single-phrase ERPs. Epochs contaminated with ocular artifacts and other artifacts were excluded from averaging on the basis of voltage values ranging from −70 to 70 μV. For baseline correction, the waves were aligned with mean voltage values in the time range between 200 msec prestimulus and the onset of the stimulus. Trials of correct responses in the comprehension task were submitted for averaging. The accepted rate for each condition (CC, MSC, LSC) per phrase was 71%, 71%, 72% for the NP1; 76%, 78%, 75% for the NP2; 75%, 75%, 74% for the NP3; 65%, 61%, 64% for VP1; 62%, 64%, 63% for the VP2; and 69%, 68%, 68% for the multiword ERPs. One-way repeated measures ANOVAs with the factor sentence type were conducted for each phrase, showing that the significant main effect of sentence type did not appear in any comparison [NP1: $F(2,26) = 0.233$, $p = .790$; NP2: $F(2,26) = 1.184$, $p = .322$; NP3: $F(2,26) = 0.128$, $p = .881$; VP1: $F(2,26) = 2.130$, $p = .139$; VP2: $F(2,26) = 0.508$, $p = .608$; multiword: $F(2,26) = 0.511$, $p = .606$]. For the purpose of presentation, grand-averaged ERP waves were smoothed using a low-pass filter with a cutoff frequency of 40 Hz. The scalp potential of a difference ERP (scrambled conditions − the CC) was reconstructed by linear interpolation (interpolation point 4) every 100-msec interval, utilizing all of the 123 scalp electrodes.

Statistical evaluation of the ERP effects was carried out separately for the multiphrase ERPs and the single-phrase ERPs for each of the three NPs and two VPs using unfiltered data of individual averaged ERPs. For an interpretation of reliable ERP effects in the scrambled conditions, the CC was used as a baseline, relative to which the polarity of reliable effects elicited by the scrambled conditions was determined. For the statistical analyses, the electrode sites that fall under the 10/20 international system (Jasper, 1958) were divided into eight lateralized regions of interest (ROIs) (nine electrodes each for left frontal [LF] and right frontal [RF]; eight electrodes each for temporal [LT, RT], parietal [LP, RP], and occipital [LO, RO] in each of the left and right hemispheres), and two midline ROIs (four electrodes for anterior [MA] and five electrodes for posterior [MP]) (see Figure 1).

We conducted 100-msec step analyses for multiword ERPs as well as for single-word ERPs, utilizing a combined and averaged signal in each ROI. Repeated measures three-way ANOVAs were performed with sentence type (CC, MSC, LSC), hemisphere (left [L], right [R]), and area (frontal [F], temporal [T], parietal [P], occipital [O]) on the lateral sites, and repeated measures two-way ANOVAs with sentence type (CC, MSC, LSC) and anteriority (anterior [MA], posterior [MP]) on the midline sites as within-participant factors. The data acquired at the lateral and midline sites were treated separately. When there was a significant interaction between sentence type and any topographical factor, subordinate repeated measures ANOVAs were performed to reveal the main effect of sentence type on the topographical level in question. When there was a main effect on sentence type, multiple comparisons of Bonferroni were conducted to examine which sentence types contributed to the difference. For all statistical tests, a p level of .05 was considered significant. The results of the Bonferroni tests represent the corrected p values. Degrees of freedom were reduced using the Greenhouse–Geisser correction method to avoid Type I errors whenever necessary (Greenhouse & Geisser, 1959). The original degrees of freedom, together with the corrected p values, are reported. Taking advantage of the multichannel EEG system, we also independently carried out paired t tests for all of the scalp electrodes (123 channels) in time windows in which significant effects including the factor sentence type were obtained, for the purpose of investigating more detailed scalp distributions of ERP effects. Electrodes in each ROI with a constantly significant difference ($p < .05$) were only reported if t tests were performed for consecutive time windows.

RESULTS

Behavioral Data

Comprehension accuracy for each condition is 92.1% (SD = 5.6) in the CC, 89.6% (SD = 6.6) in the MSC, and 88.8% (SD = 6.7) in the LSC. The mean latency of accurate responses is 1468.0 msec (SD = 625.2) in the CC, 1434.5 msec (SD = 560.2) in the MSC, and 1557.7 msec (SD = 696.5) in the LSC. Although reaction times of three sentence types were not statistically different, repeated measures ANOVA and subsequent multiple comparisons for comprehension accuracy revealed a significant difference between the CC and the LSC [$F(2,26) = 3.976$, $p = .031$; CC vs. LSC: $p = .029$]. The difference of comprehension accuracy between the CC and the LSC indicates that the participants experienced more difficulty comprehending the sentences in the LSC.

Event-related Potentials

In single-word ERPs, no significant main or interaction effects were found in the sentence-initial adverbial/postpositional phrases, the second NPs, and the first VPs in the analysis of all the conditions. Therefore, from this point on, we report the results of the phrases where the statistical analyses reached significance, namely, in the multiword ERPs, the first NPs, the third NPs, and the second VPs.

Multiword ERPs (The Three Consecutive NPs)

The grand-average ERPs for the three consecutive NPs in all the conditions at the representative electrode for each ROI is shown in Figure 2A, and the voltage maps and t maps of the difference ERPs are aligned in Figure 2B and C. Results of the grand ANOVA with 100-msec analysis are shown in Table 2. The results revealed that sentences in the LSC elicited an anterior negativity, relative to those in the CC and the MSC, which appeared at about 300 msec after the onset of the first NP and was present until the end of the second NP. Thus, a SAN was clearly observed. Although a SAN appeared to extend to the third NP in the waveforms and topography, the results of ANOVA did not prove to be significant. The lack of statistical effects of the third NP could be due to individual variability.

Whether or not a SAN effect consists of a unitary component could be an interesting question to ask next. For example, it would be quite natural to expect that a SAN would become more pronounced as a function of time. Our results, however, go against this expectation. When we look at the waveforms and topography more closely, the SAN fluctuates in amplitude and distribution. Topography of the difference ERP (LSC-CC) shows that it can roughly be divided into three phases. The first phase could comprise of a transient negativity appearing at 300–400 msec time interval of the NP1. The waveform of the LSC arrived at the peak amplitude about 350 msec poststimulus (Global field power [GFP] peak latency: 356 msec) (Lehmann & Skrandies, 1984), and the mean amplitude was more negative in the left hemisphere than in the right [300–400 msec: L = -0.72 μV (SD = 2.28); R = 0.19 μV (SD = 2.50)], confirmed by an independent ANOVA with the two factors hemisphere and area within the single level of the LSC [Hemisphere: $F(1,13) = 6.297$, $p = .026$]. The overall comparison of the three sentence types revealed the main effects of sentence type [lateral: $F(2,26) = 5.429$, $p = .011$; midline: $F(2,26) = 5.011$, $p = .014$]. Multiple comparisons showed that the main effect was due to the significant difference between the LSC and the CC [lateral: $p = .023$; midline: $p = .025$]. Paired t tests were conducted for all scalp electrodes, showing that the negativity effect was widely distributed in the left lateral areas, particularly in the left occipito-temporal site [CC vs. LSC ($p < .05$): LF = 26, 33, 34; LT = 6–16; LP = 23, 36; LO = 21, 22, 39, 40, 46, 48; MP = 65–69].

In the second phase, a slow negative deflection of the LSC appeared at about 600 msec and prolonged until 1200 msec. This slow wave is continuous to the transient negativity in the first phase, however, unlike in the first phase, it did not yield a hemispheric difference of amplitude [Hemisphere (Max: 1100–1200 msec): $F(1,13) = 2.865$, $p = .114$; Hemisphere × Area (Max: 600–700 msec): $F(3,39) = 1.010$, $p = .399$]. The overall ANOVA across the sentence types yielded the main effects of sentence type due to the significant difference of the LSC compared with the CC and the MSC throughout this time period [Max (700–800 msec): lateral, $F(2,26) = 9.536$, $p < .001$; CC vs. LSC: $p = .004$; MSC vs. LSC: $p = .007$; midline: $F(2,26) = 13.796$, $p < .001$; CC vs. LSC: $p < .001$; MSC vs. LSC: $p < .001$]. Paired t tests were also performed throughout six consecutive time windows, indicating that the negativity effect in the second phase was mainly distributed in the anterior sites [CC vs. LSC ($p < .05$): LF = 26, 27, 32–34, 53–55; LT = 8, 12, 13; LP = 36; RF = 87; RP = 92; MA = 59, 60].

The third phase comprised a negative effect, appearing at the 1400–1800 msec time period (500–900 msec within a time period of NP2) with the left frontotemporal maximum. The left-lateralized effect was confirmed statistically. The results of ANOVA for the lateral site found significant interaction of sentence type, hemisphere, and area [Max (1400–1500 msec): $F(6,78) = 2.726$, $p = .019$]. Separate ANOVAs were conducted to see the interaction of sentence type and area as well as the main effects of sentence type in two levels of hemisphere for each time window. The results showed significant main effects of sentence type in the left hemisphere at all time windows from 1400 to 1800 msec [Max (1400–1500 msec): Sentence type, $F(2, 26) = 9.226$, $p < .001$]. This lateralized effect was due to the significant difference of the LSC, compared with the CC

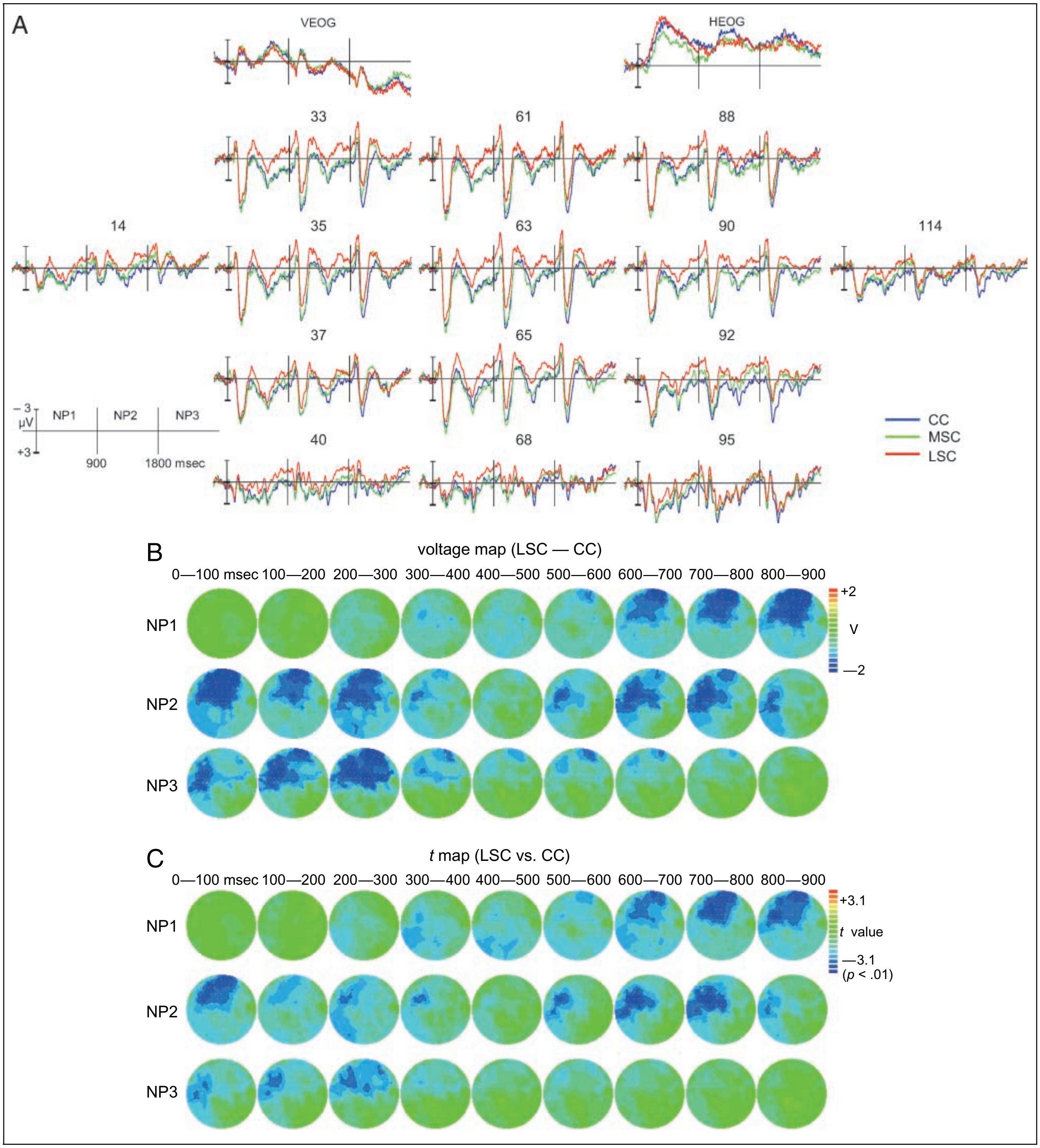


Figure 2. Grand-averaged ERP waveforms in the CC, MSC, and LSC for the multiword ERPs (the three consecutive NPs) of the representative electrode sites are superimposed in A. Negativity is plotted upward. Voltage maps of the difference ERPs (LSC − CC) and *t* maps are shown every 100 msec in B and C, respectively. Red indicates positivity and blue indicates negativity in the voltage maps and *t* maps.

and the MSC [1400–1500 msec: CC vs. LSC, $p = .006$; MSC vs. LSC, $p = .026$], which was also confirmed by paired *t* tests for individual electrodes. Electrodes with a significant difference throughout four consecutive time windows, mainly distributed at the left fronto-temporal sites, are as listed below [CC vs. LSC ($p < .05$): LF = 26, 27, 33, 36; LT = 6–8, 12, 13, 15, 16; LP = 36; LO = 21, 22].

To summarize, although the negative effect of the LSC remains constant throughout [electrodes with a constantly significant difference between the CC and the LSC in paired *t* tests for three phases: LF = 26, 33; LT = 8, 12, 13; LP = 36], the topography changes from the left hemisphere (occipito-temporal area) to the anterior area, and then back again to the left hemisphere (fronto-temporal

Table 2. *F* Values for the ANOVAs at the Three Consecutive NPs

Effects	*df*	*Time Windows of NP1 (msec)* 0–100	100–200	200–300	300–400	400–500	500–600	600–700	700–800	800–900
Lateral										
Type	(2,26)				5.429*	3.496*		6.822**	9.536**	5.172*
Type × Hemisphere	(2,26)									
Type × Area	(6,78)									
Type × Hemisphere × Area	(6,78)									
Midline										
Type	(2,26)				5.011*	4.290*	5.847**	11.038***	13.796***	9.130**
Type × Area	(2,26)									

Effects	*df*	*Time Windows of NP2 (msec)* 900–1000	1000–1100	1100–1200	1200–1300	1300–1400	1400–1500	1500–1600	1600–1700	1700–1800
Lateral										
Type	(2,26)	4.300*	3.544*	6.297**	4.942*		6.344**	5.268*	4.514*	
Type × Hemisphere	(2,26)							3.452*		
Type × Area	(6,78)									
Type × Hemisphere × Area	(6,78)						2.726*	2.281*	2.265*	2.478*
Midline										
Type	(2,26)	6.484**	4.945**	5.166*	4.638*		4.483*	3.942*		
Type × Area	(2,26)									

Effects	*df*	*Time Windows of NP3 (msec)* 1800–1900	1900–2000	2000–2100	2100–2200	2200–2300	2300–2400	2400–2500	2500–2600	2600–2700
Lateral										
Type	(2,26)									
Type × Hemisphere	(2,26)									
Type × Area	(6,78)									
Type × Hemisphere × Area	(6,78)									
Midline										
Type	(2,26)									
Type × Area	(2,26)									

Type = sentence type.

*$p < .05$.

**$p < .01$.

***$p < .001$.

area) as a function of time. These results suggest that the observed negativity does not seem to be a unitary component and has multiple physiological bases, and thus, involves multiple cognitive functions.

ERPs for the First NP

The grand-average ERPs for the three sentence types are plotted and superimposed in Figure 3A, and the voltage maps and t maps of the difference ERPs (LSC − CC) are aligned in Figure 3B and C, respectively. The overall pattern of waveforms and topography for the first NP by the single-word analyses resembles those of the corresponding time period of the multiword ERPs. At the early time range (200–500 msec), corresponding to the first phase in the multiword ERPs, the negative effect of the LSC was observed through three consecutive time windows [Max (300–400 msec): lateral, sentence type, $F(2,26) = 6.237$, $p = .006$; midline: sentence type, $F(2,26) = 5.128$, $p = .013$] (Table 3). Subsequent pairwise comparisons revealed that the differences between the LSC and the other two types were significant at two time windows from 300 to 500 msec [300–400 msec: lateral, CC vs. LSC, $p = .019$; MSC vs. LSC, $p = .047$; midline: CC vs. LSC, $p = .042$]. Paired t tests utilizing individual electrodes at the two consecutive intervals from 200 to 400 msec showed that the similar negative effect as in the first phase of the multiword ERPs emerged, in particular, at the left lateral areas [CC vs. LSC ($p < .05$): LF = 26, 32, 33; LT = 6, 12, 13, 15; LO = 21, 22, 40, 41]. Due to the lack of interaction of sentence type and topographic factors in the grand ANOVA, it could not be strongly claimed that the early negative effect has a left dominant distribution, however, potential distribution of this early negativity in the left hemisphere is suggested.

At the later time interval of 500–900 msec, almost parallel to the second phase in the multiword ERP, the similar negative effect in anterior sites was observed in the single ERP analysis. At the 500–600 msec time interval, the topography shortly changed from the previous time windows, and a negative effect appeared in

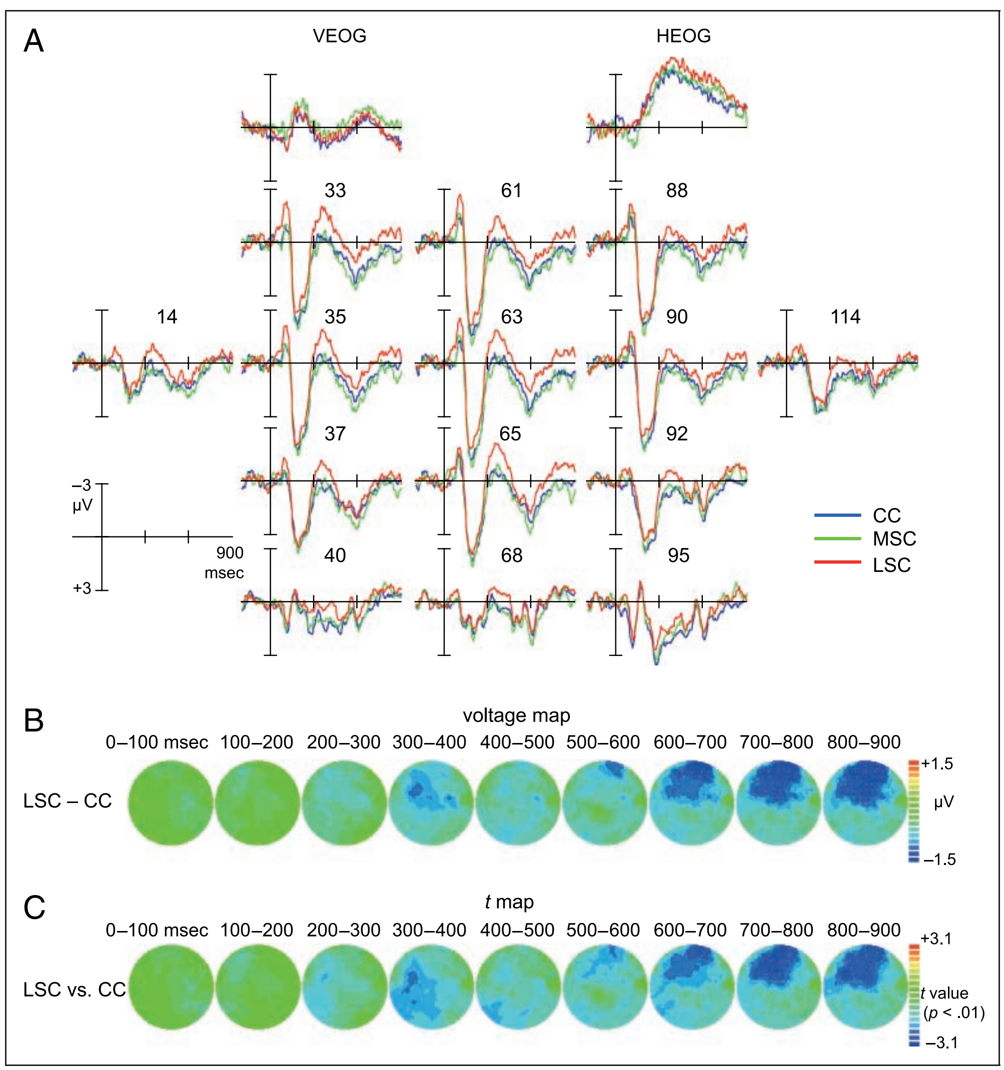


Figure 3. Grand-averaged ERP waveforms in the CC, MSC, and LSC for the first NP of the representative electrode sites are superimposed in A. Negativity is plotted upward. Voltage maps of the difference ERPs (LSC − CC) and t maps are shown every 100 msec in B and C, respectively.

Table 3. *F* Values for the ANOVA at the First NP

Effects	*df*	*Time Windows (msec)*								
		0–100	*100–200*	*200–300*	*300–400*	*400–500*	*500–600*	*600–700*	*700–800*	*800–900*
Lateral										
Type	(2,26)			3.574*	6.237**	4.415*	5.282*	6.961**	6.245**	6.691**
Type × Hemisphere	(2,26)									
Type × Area	(6,78)								4.066*	
Type × Hemisphere × Area	(6,78)									
Midline										
Type	(2,26)				5.128*	3.928*	6.359**	7.170**	8.791**	9.098**
Type × Area	(2,26)								4.622*	3.388*

Type = sentence type.

*$p < .05$.

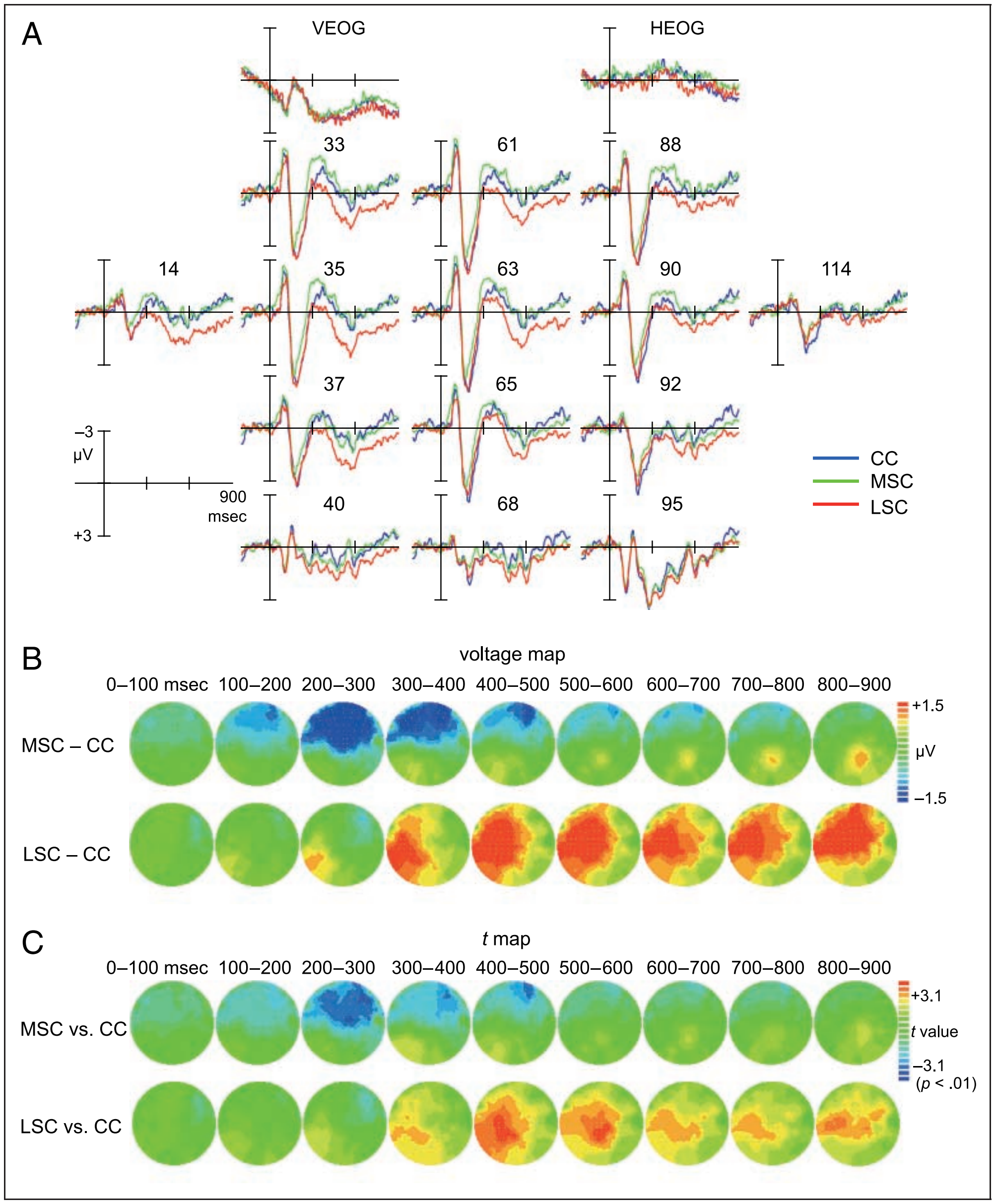
**$p < .01$.

the anterior region and became more prominent bilaterally as a function of time (Figure 3B and C). This anterior negativity is confirmed by an ANOVA. Significant interaction of sentence type and area was found from the 700 msec onward. At 700–800 msec, the interaction Sentence type × Area reached a significant level in both the lateral and midline sites [lateral: $F(6,78) = 4.066$, $p = .014$; midline: $F(2,26) = 4.622$, $p = .036$]. Separate ANOVA, followed by multiple comparison for each area, revealed a significant difference between the CC and the LSC in the three lateral and midline anterior sites [F: Sentence type, $F(2,26) = 12.052$, $p < .001$; CC vs. LSC, $p < .001$; T: Sentence type, $F(2,26) = 5.640$, $p = .009$; CC vs. LSC, $p < .001$; P: Sentence type, $F(2,26) = 3.577$, $p = .042$; CC vs. LSC, $p < .001$; MA: Sentence type, $F(2,26) = 13.157$, $p < .001$; CC vs. LSC, $p < .001$]. At the subsequent time interval (800–900 msec), the interaction of sentence type and area was shown in the midline [$F(2,26) = 3.388$, $p = .049$]. The effects of sentence type proved to be significant both in the anterior and the posterior sites [MA: $F(2,26) = 10.465$, $p < .001$; MP: $F(2,26) = 3.808$, $p = .035$], however, the significant difference between the CC and the LSC was only observed in the anterior ROI [MA: $p < .001$; MP: $p = .091$]. Comparisons between the CC and the LSC for single electrodes also confirmed the anterior distribution of the negativity effect throughout four consecutive time windows [CC vs. LSC ($p < .05$): LF = 26, 32–34, 53–55; LT = 12, 13; RF = 82, 83, 87, 107; MA = 59–62].

ERPs for the Third NP

The grand-average ERPs, the voltage maps, and *t* maps of the difference ERPs of the third NP are aligned in Figure 4A, B, and C, respectively. The ERP waveform of the LSC deflected in the positive direction relative to that in the CC after 300 msec post-stimulus, peaking around 550 msec (GFP peak latency: 544 msec). This positive wave had a hemispheric difference of amplitude in the frontal and temporal ROIs during 400–600 msec [500–600 msec: LF = 1.87 μV (*SD* = 3.18); RF = 1.03 μV (*SD* = 2.64); LT = 1.94 μV (*SD* = 2.51); RT = 0.80 μV (*SD* = 1.60); Hemisphere × Area: $F(3,39) = 3.179$, $p = .035$; Hemisphere: LF, $F(1,13) = 6.661$, $p = .023$; LT, $F(1,13) = 8.313$, $p = .013$]. This positive component is likely to be a P600. At the 300–400 msec time interval, a positive effect of the LSC appeared in the left temporo-parietal regions. The statistical analyses confirm this observation (Table 4). The three-way interaction of sentence type, hemisphere, and area proved to be significant [$F(6,78) = 3.128$, $p = .035$]. The interaction of sentence type and area for each hemisphere also reached significance [L: $F(6,78) = 7.943$, $p < .001$; R: $F(6,78) = 6.441$, $p = .002$]. Further analyses revealed that although there were significant effects of sentence type for all areas in both hemispheres, the difference between the CC and the LSC was significant in three areas of the left hemisphere [LT: $p = .036$; LP: $p = .048$; LO: $p = .016$]. The positivity becomes larger in magnitude and broader in distribution during the next 200-msec interval. This observation was confirmed statistically. At the two time windows from 400 to 600 msec, the interactions Sentence type × Hemisphere and Sentence type × Area proved to be significant [400–500 msec: lateral, Sentence type × Hemisphere, $F(2,26) = 4.102$, $p = .028$; Sentence type × Area, $F(6,78) = 5.596$, $p = .003$; midline: Sentence type × Area, $F(2,26) = 5.733$, $p = .009$]. Pairwise comparisons, following tests of the

Figure 4. Grand-averaged ERP waveforms in the CC, MSC, and LSC for the third NP of the representative electrode sites are superimposed in A. Negativity is plotted upward. Voltage maps of the difference ERPs (MSC − CC, and LSC − CC) and *t* maps are shown every 100 msec in B and C, respectively.



effect of sentence type for each topographical level, revealed that the difference between the CC and the LSC emerged over all scalp areas [lateral: L, $p < .001$; R, $p = .002$; F, T, P, O, $p < .001$; midline: MA, $p = .002$; MP, $p = .033$]. This strong positivity still continues to appear at subsequent time windows. At the 700–800 msec time interval, Sentence type × Area interaction reached a significant level [lateral: $F(6,78) = 4.070$, $p = .017$; midline: $F(2,26) = 6.063$, $p = .007$]. Further analyses showed that the main effects of sentence type were significant at three lateral (F, T, P) and midline anterior sites [F: $F(2,26) = 5.876$, $p = .008$; T: $F(2,26) = 6.740$, $p = .004$; P: $F(2,26) = 4.476$, $p = .021$; MA: $F(2,26) = 5.443$, $p = .022$], and a highly significant difference was revealed between the CC and the LSC in the three lateral areas ($p < .001$). Paired *t* tests throughout six continuous intervals from 300 to 900 msec also confirmed the persistent positive effect of the LSC, especially in the left temporal ROI [CC vs. LSC ($p < .05$): LT = 6–8, 13–16; LP = 23, 24, 38; LO = 22, 39, 48; MP = 67].

In addition, we found that sentences in the MSC elicited a transient negativity, relative to those in the CC, at an early time window of 200–300 msec, 100 msec prior to the appearance of the positive effect of the LSC, in the right anterior region. The interaction of sentence type and topographic factors proved to be significant [lateral: Sentence type × Hemisphere × Area, $F(6,78) = 3.294$, $p = .027$; midline: Sentence type × Area, $F(2,26) = 5.132$, $p = .013$]. Subsequent analyses following ANOVAs for each hemispheric level and each midline site confirmed that the difference between the CC and the MSC was only significant in the right frontal, temporal, and parietal areas, as well as in the midline

Table 4. *F* Values for the ANOVA at the Third NP

Effects	*df*	*Time Window (msec)* 0–100	100–200	200–300	300–400	400–500	500–600	600–700	700–800	800–900
Lateral										
Type	(2,26)			7.299**	10.057***	13.833***	9.047**	6.291**	5.223*	6.398**
Type × Hemisphere	(2,26)			4.626*		4.102*	3.695*			
Type × Area	(6,78)			7.367**	7.927***	5.596**	3.223*		4.070*	
Type × Hemisphere × Area	(6,78)			3.294*	3.128*					
Midline										
Type	(2,26)		3.741*	8.425**	8.535**	11.138***	7.400**	4.298*	3.649*	5.503*
Type × Area	(2,26)			5.132*	5.123*	5.733**	4.397*	5.533*	6.063**	4.328*

Type = sentence type.

*$p < .05$.

**$p < .01$.

***$p < .001$.

anterior site [RF: Sentence type, $F(2,26) = 13.054$, $p < .001$; CC vs. MSC, $p = .002$; RT: Sentence type, $F(2,26) = 8.092$, $p = .002$; CC vs. MSC, $p = .004$; RP: Sentence type, $F(2,26) = 6.892$, $p = .004$; CC vs. MSC, $p = .017$; MA: Sentence type, $F(2,26) = 10.269$, $p < .001$; CC vs. MSC, $p = .008$]. The significant three-way interaction for the lateral site was also observed at the next time window of 300–400 msec, and follow-up analyses showed that the difference between the CC and the MSC remained significant in the right frontal and temporal ROIs [RF: Sentence type, $F(2,26) = 13.288$, $p < .001$; CC vs. MSC, $p = .008$; RT: Sentence type, $F(2,26) = 9.409$, $p < .001$; CC vs. MSC, $p = .029$]. At the 400–500 msec interval, the negative effect was still persistent in anterior portions [F: Sentence type, $F(2,26) = 15.555$, $p < .001$; CC vs. MSC, $p < .001$; MA: Sentence type, $F(2,26) = 15.073$, $p < .001$; CC vs. MSC, $p = .007$]. A paired *t* test for each electrode at three consecutive time windows from 200 to 500 msec warranted the right anterior negative effect of the MSC [CC vs. MSC ($p < .05$): RF, 83, 87, 88; MA, 60].

ERPs for the Second VP

The grand-average ERPs and the topography for the second VP showed that sentences in the scrambled conditions (LSC, MSC) elicited an anterior negativity, relative to those of the CC, at the time window between 200 and 600 msec (Figure 5A, B, and C). This anterior negativity was reflected in the significant interaction of sentence type and area at the three consecutive time windows from 200 to 500 msec interval [Max (200–300 msec): lateral, $F(6,78) = 6.651$, $p < .001$; midline: $F(2,26) = 4.883$, $p = .016$] (Table 5). A separate ANOVA revealed that the effects of sentence type were significant at the more anterior sites [200–300 msec: F, $F(2,26) = 13.572$, $p < .001$; T, $F(2,26) = 5.867$, $p = .008$; P, $F(2,26) = 5.977$, $p = .007$; MA, $F(2,26) = 15.974$, $p < .001$] due to the significant difference of the LSC and the MSC, compared with the CC [CC vs. MSC: F, $p < .001$; P, $p = .006$; MA, $p < .001$; CC vs. LSC: F, $p < .001$; P, $p = .002$; MA, $p = .001$]. These negative effects of the LSC and the MSC were confirmed by paired *t* tests for individual electrodes [CC vs. LSC ($p < .05$): LF = 55; RF = 81–83, 88, 89, 106; RP = 79, 91; MA = 59–62; CC vs. MSC ($p < .05$): RF = 81; RP = 79, 91; MA = 59]. The negative effect of the LSC remained at the 500–600 msec interval [lateral: Sentence type, $F(2,26) = 8.617$, $p = .001$; CC vs. LSC, $p = .003$; midline: Sentence type, $F(2,26) = 7.875$, $p = .002$; CC vs. LSC, $p = .005$], suggesting that the processing of the LSC takes approximately 100 msec longer than the MSC.

DISCUSSION

The present study set out to investigate whether a SAN was elicited by the NP-movement and to differentiate anterior negativities into subcomponents. To summarize the results, the SAN was modulated at the first two NPs of the sentences in the LSC, relative to the CC. Subsequently, a widely distributed positive component was observed after the 300 msec at the pregap position (NP3) in the LSC. At the same position in the MSC, an early negative component appeared at the 200–500 msec time interval. Finally, at the second VP, a bilateral negative component was observed in the anterior electrodes in both the MSC and the LSC, relative to the CC.

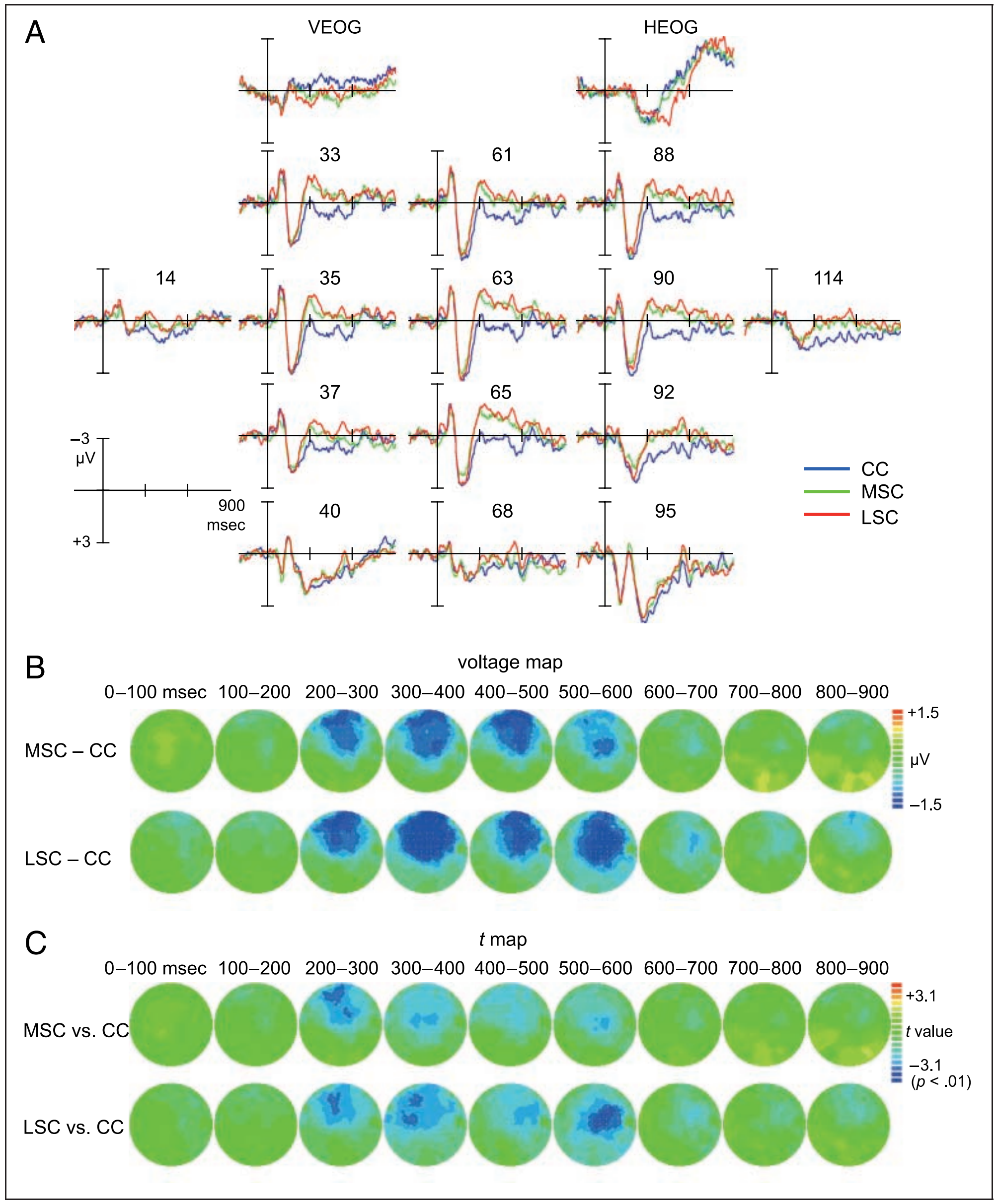


Figure 5. Grand-averaged ERP waveforms in the CC, MSC, and LSC for the second VP of the representative electrode sites are superimposed in A. Negativity is plotted upward. Voltage maps of the difference ERPs (MSC − CC, and LSC − CC) and *t* maps are shown every 100 msec in B and C, respectively.

The prediction that more complex constructions of the scrambled word order in the NP-filler–gap dependency would elicit a sustained negative effect was borne out. The SAN in the LSC, together with a lack of it in the MSC, thus confirms the *Complexity Hypothesis* of the scrambled word order in relation to the storage cost. On the other hand, that the differences in the property of the *wh*-filler and the NP-filler would contribute to the SAN effect was not borne out, hence, the *Displacement Type Hypothesis* was not supported as far as the SAN is concerned. In addition, a P600 was observed at the third NP of the pregap position, suggesting that the *Integration Hypothesis* was supported. However, the position in the sentence in which the P600 appeared makes one wonder if the functional nature of it differs from the notion of *Integration* that is reported in the ERP literature in English (Phillips et al., 2005; Kaan et al., 2000). The order of the remainder of this section is as follows. First,the SAN in the multiphrase ERPs will be discussed. We will then focus on the P600 in the third NP, followed by the anterior negativity in the second VP.

A Sustained Anterior Negativity Reflecting Storage Cost in NP-Movement

The finding of a SAN spanning over multiphrases is in line with previous studies on filler–gap dependencies of the *wh*-type (e.g., Ueno & Kluender, 2003; Fiebach et al., 2002; King & Kutas, 1995; Kluender & Kutas, 1993). Furthermore, our results are compatible with earlier studies that reported a SAN in NP-movement (Matzke et al., 2002). The topographical distribution of the left hemisphere negativity matches that reported by Fiebach et al. (2002), Matzke et al. (2002), and King and Kutas

Table 5. *F* Values for the ANOVA at the Second VP

		Time Window (msec)								
Effects	*df*	*0–100*	*100–200*	*200–300*	*300–400*	*400–500*	*500–600*	*600–700*	*700–800*	*800–900*
Lateral										
Type	(2,26)			7.373**	8.050**	4.185*	8.617**			
Type × Hemisphere	(2,26)							3.940*		
Type × Area	(6,78)			6.651***	4.936**	3.578*				
Type × Hemisphere × Area	(6,78)									
Midline										
Type	(2,26)			8.568**	6.509**	4.248*	7.875**			
Type × Area	(2,26)			4.883*						

Type = sentence type.

*$p < .05$.

**$p < .01$.

***$p < .001$.

(1995). Accordingly, our findings show that NP scrambling in complex sentences creates a filler–gap dependency, and thus, holding an NP-filler while analyzing syntactic relations of incoming stimuli results in storage cost of working memory in sentence comprehension.

There are several important issues to be mentioned. First, the sustained negativity in our study lasted until the end of the NP2 and did not continue to the NP3 of the pregap position, whereas in other studies the sustained effect extended to the pregap element. The absence of a SAN at the NP3 may be attributed to the prominent positive effect elicited at the NP3, which may have canceled out the sustained negative effect. Second, although our findings are in accord with previous studies of the NP-filler–gap studies of scrambling in German reported by Matzke et al. (2002), showing a deflection starting at the accusative-marked article in sentence-initial position, they differ from the *wh*-filler–gap dependency reported by Fiebach et al. (2002), Kluender and Kutas (1993), and King and Kutas (1995). In the *wh*-filler–gap studies, the negative deflection started at the words following the *wh*-phrase. In other words, no effects were observed at the *wh*-filler itself. The reason for this difference in the onset of the deflection might depend on the type of filler: An NP-filler itself may be easier to detect and may be more prone to eliciting a sustained negativity than a *wh*-filler, but to confirm this possibility, more work will have to be done in Japanese with sentences of scrambled word order using both an NP-filler and a *wh*-filler.

Third, the tripartite nature of the SAN could be accounted for in the following way. In the first phase, a transient negativity appeared at the 300–400 msec time interval of NP1 in the LSC, compared to the CC, in which the mean amplitude was more negative in the left hemisphere than in the right hemisphere. The same negative effect was also confirmed by the single-word ERP analyses of the NP1. This finding is consistent with that reported by Schlesewsky et al. (2003), Matzke et al. (2002), and Rösler et al. (1998): The latency of the negativity started as early as 200 msec and lasted for 200 msec. Based on the fact that a transient LAN has been ubiquitously reported in ungrammatical sentences with various morphosyntactic violations (Friederici, Pfeifer, & Hahne, 1993; Rösler, Puetz, Friederici, & Hahne, 1993; Neville, Nicol, Barss, Foster, & Garrett, 1991; Kutas & Hillyard, 1983), Friederici et al. (2003) claimed that scrambled word order in grammatically *congruent* sentences induces a syntactic violation effect due to a mismatch of the predicted structural position and the input item. In this respect, our study on the scrambling of an NP, involving complex sentences of a case-marking language with a highly flexible word order, is consistent with this view. The parser incrementally predicts what the next structural position will be on the basis of incoming elements and the structure built so far. When the parser encounters an object NP with the accusative marker *–o* of a scrambled word order sentence in the LSC, the only available structural position is the specifier of the topic phrase, the position for the topic NP with the topic marker *–wa*. Here the parser recognizes the mismatch between the object NP with the accusative marker *–o* and the position available to it. This structural mismatch may have induced the transient LAN, reflecting a syntactic violation.

On the other hand, the transient negative effect in the first phase has distinct characteristics from those reported in previous related studies. The topographical distribution of this component (i.e., the negativity of the left occipito-temporal maximum) slightly differs from the typical anterior distribution of the LAN for phrase structure violations, as reported in the literature about German. The left posterior distribution may have originated from the internal organization of words and phrases in Japanese, namely, how grammatical case is marked with respect to sentence structure. In Japanese, grammatical case is expressed by a case particle, a bound morpheme which functions like a suffix, following the noun that it attaches to (e.g., *Ken–ga/–o/–ni/–no*). If we suppose that the scanning of words and phrases is performed in a manner from left to right, the parser would first encounter a noun in Japanese, and not avoid processing the interpretive properties of the noun at the same time it checked grammatical case. In German, on the other hand, grammatical case is encoded in an article, a free morpheme, always preceding a noun (e.g., *der/den/dem/des Mann(s)*). If the parser processes an article before processing a noun in German, syntactic information would be processed more prominently than the semantic information carried by the noun. Assuming a widely known observation that semantic processing is subserved by the left temporal region, whereas syntactic information is carried out in a lower part of the left frontal area, it would be possible that the processing order of a noun and a case marker modulates the scalp distribution of the left negativity.

A similar left posterior negative effect was reported in processing *wh–wh* ungrammatical sentences in English (e.g., ''What did he wonder who he could coerce INTO?'') in Kluender and Kutas (1993). However, the eliciting word was not the object *wh*-phrase in the sentence initial position but the preposition immediately following the embedded direct object gap. In the present study and the studies of Kluender and Kutas, the critical word is a content word that contains a certain amount of semantic information: It is an ordinary NP in our study, and it is a preposition linked with the displaced *wh*-phrase in the study of Kluender and Kutas. It is possible that the object-related element temporarily yielded the left hemisphere negativity and that the first processing of content words induced more posterior distribution. The discrepancy of the position yielding the left posterior negativity between both studies, on the other hand, would come from the difference of the critical point at which the grammatical status of the displaced constituent is detected. In our study, the object NP is identified and elicits the negativity effect in the displaced position because the accusative case marker *–o*, which denotes the noncanonicity of the word order, attaches to it. In the Kluender and Kutas' study, the grammatical role and ungrammatical status of the displaced *wh*-phrase were finally confirmed at the location of the preposition for the reason that the *wh*-phrase was not assigned an overt morphological case marker, not obviating the grammatical status until the presence of the preposition.

To summarize, it is suggested that the occipito-temporal distribution of the left hemisphere negativity in the first phase results from the interaction of several factors, including the order of nouns and case markers, the detection timing of grammatical status, and semantic properties of displaced constituents.

Turning back to the second phase of the SAN, at the 600–1200 msec time interval, the anterior negativity involves the frontal electrodes. Given that this is the period that the parser maintains the object NP with accusative marker *–o* in working memory while encountering and analyzing the second NP with the topic marker *–wa*, the negativity of time interval would be the genuine component of the SAN, reflecting the function of storing a filler in working memory. In fact, the timing of 600 msec coincides with the offset time when the stimuli disappeared and the parser actually started to memorize and store the first NP in working memory. The third phase of the 1400–1800 msec interval is also an extension of the same function: The parser maintains two NPs, *NP–o* (accusative) and *NP–wa* (topic), in working memory. At the same time, it anticipated what the next element would be (e.g., a noun or a verb) and prepared for the structural analysis of the sentence. Here again a kind of linguistic analysis and processing is required more so than in the previous phase, hence, resulting in more involvement of the left hemisphere.

A P600 at the Preverbal Position Reflecting a Structural Integration

The positivity observed at the third NP of the pregap position after about 300 msec poststimulus, peaking around 600 msec, resembles a typical P600 component in morphology. The topographical distribution, however, has a left fronto-temporal maximum, different from those of P600s in previous studies. The distribution of the P600 in our study is partly in line with previous findings of the P600 with the relatively anterior maximum in the processing nonpreferred continuation of the preceding fragments in grammatically correct sentences, and the revision process, as opposed to the posteriorly distributed P600, reflecting the repair process in ungrammatical sentences (Hagoort, Brown, & Osterhout, 1999; but see Kaan & Swaab, 2003 for a different interpretation). However, the cognitive processes involved in the sentences of the LSC are different from the revision process found in nonpreferred ambiguous sentences. The sentences in the LSC are not ambiguous, hence, the third NP is not the point at which the ambiguity resolution occurs.

Other studies reported that syntactically complex sentences yielded a frontal P600 (Kaan & Swaab, 2003; Friederici, Hahne, & Saddy, 2002). Although scrambled word order with long-distance filler–gap dependency is far more complex than canonical word order without such a dependency, as shown in the behavioral accuracy data, the left lateralized P600 in our study is different from a frontal P600 in the scalp distribution. We suggest that the P600 with a left fronto-temporal maximum is a reflection of the cost of structural integration, a subtype of the syntactic integration difficulty of Gibson (1998, 2000) and Kaan et al. (2000). Because the P600 was elicited before the input of the verb in our study, the cognitive processes at issue are naturally different from those of the filler–gap dependency in English *wh*-questions where the P600 occurred at the verb itself (Phillips et al., 2005; Kaan et al., 2000). That is, Phillips et al.'s interpretation of the P600 as syntactic confirmation of thematic role assignment and semantic interpretation of the verb and the filler is not suitable to explain the pregap positivity in verb-final languages. In our study, the parser has to construct a structural tree solely on the basis of syntactic information, that is, the order of case particles, because the parser is required to process the third NP without the semantic cue of the related verb. Additionally, because the nouns used in our study are all animate, and the thematic roles are all reversible (a noun can be the agent or patient of the embedded verb or the agent of the matrix verb), as stated in the Methods section, the properties of the nouns used cannot assist in the compositional interpretation of the sentences. That is, the parser has to try to construct structures solely on the basis of the order of case particles.

Our claim that the left fronto-temporal P600 reflects the structural integration cost could be compatible with the dissociation of the distribution of the P600 between the present and those investigating the dependency-length effect of the *wh*-filler and its gap in English (Phillips et al., 2005). The P600 in Phillips et al. had been primarily observed at the posterior electrodes for both the short and long dependency conditions. The discrepancy between the topographical distributions in our study and those in Phillips et al. appears to be due to the difference in word order between English and Japanese. The process of thematic role assignment and the subsequent compositional interpretation play a role in their study because the critical word in English is a verb, whereas no such semantic processes come into play in our study of Japanese, where the verb appeared after the critical word. The posterior P600 is also reported in studies utilizing sentences with the collapse of the structural representation that requires to be repaired (Kaan & Swaab, 2003; Hagoort et al., 1999). Because the stimuli sentences we used are structurally congruent, our results are not incompatible with the previous observation about the posterior P600 as a reflection of the repairing process.

The plausibility of this hypothesis about the left lateralized P600, furthermore, comes from the ERP component observed at the same position in the MSC. At the third NP of the MSC, an early negativity was elicited at the TW of 200–500 msec but no P600 was observed. Assuming that the order of *NP-wa* (topic) *NP-o* (accusative) is canonical in Japanese, the input *NP-wa* followed by *NP-o* in the MSC makes the parser expect a transitive verb to appear next, more so than the *NP-ga* (nominative). This is not the case in the LSC because at the first NP (*NP-o*) the parser already notices that the word order is non-canonical. The order of *NP-o NP-wa* makes it extremely difficult for the parser to make any decisive prediction regarding the next input at this point, more so than in the CC and MSC. It is at this place in the LSC that a great amount of processing resources are consumed in the competition of several possible inputs for constructing the most appropriate structural tree.

To summarize, converging evidence of the P600 in the LSC, together with the early negativity at the 200–300 msec interval and the absence of the P600 in the MSC, suggests that our interpretation of the pregap positivity as the cost of structural integration based on case information in congruous sentences is on the right track.

Recomputation of Scrambled Word Order at the Matrix Verb Phrase

One of the most novel findings in our study was the frontally distributed negativity observed at the second VP, the matrix verb, which was in sharp contrast with the absence of ERP effects at the first VP, the embedded verb. This negativity was elicited in both the MSC and the LSC. In the stimuli sentences used in our study, the checking and confirmation of grammatical relations in all the NPs must take place at the matrix verb, the second VP, crucially *not* at the first VP of the embedded clause. Here, the parser must first check the subject and the object of the complement clause, the subject of the main clause, and then identify which NP is the agent and the patient of the complement verb and which NP is the agent of the main verb in order to understand the meaning of the sentence. The compositional interpretation of sentences requires proper assignment of thematic roles to NPs. It is at the matrix verb that the number of arguments in the argument structure of the predicates (agent and patient with transitive verbs in the embedded clause and agent of the matrix verb in the main clause) must match with the number of NPs processed thus far. At the first VP, only two of the thematic roles are provided, whereas there are three NPs that thematic roles need to be assigned to. Because thematic role assignment to the preceding NPs cannot be completed at the position of the embedded verb, the first VP, the parser has to prolong this operation until the second VP, where all the thematic roles will be assigned. The absence of any ERP effects at the first VP shows that this was, in fact, the case.

Here we suggest that the negativity observed here is a reflection of the recomputational operations. Considering that the structure-building of the preceding three NPs has already been performed at the third NP, as suggested by the P600 in the LSC, what was left to be done here is the rechecking of the grammatical relations of the embedded clause and matrix clause, the syntactic assignment of thematic roles to arguments along with its reconfirmation, and finally, the compositional interpretation of the entire sentence. These operations are equally necessary for all the experimental sentences in our stimuli regardless of sentence types. No other factors are involved that could cause the difference in ERPs. Hence, the observed ERP component at the matrix VP is most likely a reflection of the cost specifically required to process scrambled word order. Reordering of the three NPs and recomputing the grammatical relations and thematic roles of the two clauses, which cause an additional cost in processing, are not necessary operations for the nonscrambled canonical sentences. The fact that the topography of voltage maps and *t* maps in the MSC and those of the LSC, relative to the CC, resemble one another provides further support for this claim.

Although previous studies tried to interpret this negativity at the postgap position as ''the overall difficulty of processing containing filler–gap dependencies'' (Kluender & Kutas, 1993: 263), ''the difficulty in thematic role assignments'' (King & Kutas, 1995: 338), and ''the global working memory demands of having had to processes a noncanonical scrambled structure'' (Ueno & Kluender, 2003: 266), the exact nature of this effect was not clear due to the absence of the P600 at the pregap position. Because we used precisely the same number of words with the same array, NPs of similar properties and multiple thematic roles in different types of scrambled word order in complex sentences with a high-density EEG system, our results are the first to demonstrate the authenticity of this component in filler–gap dependencies. The combination of the P600 at the pregap position and the anterior negativity at the postgap verbal position provides a more precise explanation of this negativity than the previous studies ever have.

Concerning the topographical distribution, bilateral anterior negativity makes one wonder if nonlinguistic factors of some kind are involved in addition to the linguistic recomputational operations of the kind mentioned above. There are several possible interpretations for this activation.[2] They come from previous neuroimaging studies using functional magnetic resonance imaging. Grossman, Crino, Reivich, Stern, and Hurtig (1992) investigated the processing of sentences with relative clauses by Parkinson's patients and found activation not only of the dorsolateral prefrontal cortex but also of the left anterior cingulate. The former is related to working memory functions and the latter plays an important role in the use of attentional resources (Petrides, Alivisatos, Meyer, & Evans, 1993). More recently, Botvinick, Nystrom, Fissell, Carter, and Cohen (1999) showed that the anterior cingulate cortex of the right hemisphere is involved in performance monitoring, namely, when engaged in conflicting or competitive responses (see also Barch, Braver, Sabb, & Noll, 2000). Recomputational operations, such as the reordering of NPs, the confirmation of grammatical relations, and the assignment of thematic roles, certainly require a great amount of resources in working memory and much attention. These operations are also equally likely to be accompanied by monitoring processes. Bilateral anterior activation at the 200–500 msec interval suggests that these functionally different operations take place within the same time frame.

Conclusion

While examining a number of ERP effects on the processing of scrambled word order in Japanese, we have shown that the parser is performing quite distinctive operations phrase by phrase, each of which are subserved by different neural circuitries. Our results of the Japanese filler–gap dependency of the NP-type are the first to functionally differentiate structural integration from thematic role assignment and compositional interpretation of sentential meaning by showing distinct ERP components: The former is reflected in the P600 at the pregap position and the latter in the anterior negativity at the postgap position. The present study has demonstrated that evidence from Japanese, a language structurally different from English and German, can not only be used to refine the intricate physiological nature of processing filler–gap dependencies but also to uncover the universal and language-specific mechanisms of processing displacement phenomena.

Acknowledgments

We thank Kazuyuki Nakagome, Kenji Itoh, and Shinya Kuriki for their helpful comments and technical advice on the ERP analyses. The research presented here was supported by a Grant-in-Aid for Scientific Research from the Japanese Ministry of Education, Culture, Sports, Science and Technology (MEXT): (A)15202011, (B)15320053, and (B)15300148 to the first author.

Reprint requests should be sent to Hiroko Hagiwara, Department of Language Sciences, Graduate School of Humanities, Tokyo Metropolitan University, Minami Osawa, Hachioji, Tokyo, 192-0397, Japan or via e-mail: hagiwara@bcomp.metro-u.ac.jp.

Notes

1. Of note is that in topicalization, an instance of NP-movement, Felser et al. (2003) reported a SAN effect prior to the processing of the verb in sentences of an object-topicalized construction.
2. Needless to say, this negativity obviously differs from brain potentials of the contingent negative variation (CNV), which are frequently elicited in an unexpected event by a prime—the associated target stimulus, in waveforms and topographical distribution. Although the CNV has a slow negative wave often maximal over the fronto-central area, the negativity in the second VP is transient. Moreover, in the current study, it is unlikely that the anterior negativity is a mere sentence wrap-up

effect as seen in the previous studies which reported an N400 difference (Hagoort, Brown, & Groothusen, 1993; Osterhout & Holcomb, 1992). Neither the morphology nor the topographical distributions of our study resemble those of the N400.

REFERENCES

Barch, D. M., Braver, T. S., Sabb, F. W., & Noll, D. C. (2000). Anterior cingulated and the monitoring of response conflict: Evidence from an fMRI study of overt verb generation. *Journal of Cognitive Neuroscience, 12,* 298–309.

Bornkessel, I. D., Fiebach, C. J., & Friederici, A. D. (2004). On the cost of syntactic ambiguity in human language comprehension: An individual differences approach. *Cognitive Brain Research, 21,* 11–21.

Botvinick, M., Nystrom, L. E., Fissell, K., Carter, C. S., & Cohen, J. D. (1999). Conflict monitoring versus selection-for-action in anterior cingulated cortex. *Nature, 402,* 179–181.

Coulson, S., King, J., & Kutas, M. (1998). Expect the unexpected: Event-related brain responses of morphosyntactic violations. *Language and Cognitive Processes, 13,* 21–58.

Felser, C., Clahsen, H., & Münte, T. F. (2003). Storage and integration in the processing of filler–gap dependencies: An ERP study of topicalization and wh-movement in German. *Brain and Language, 87,* 345–354.

Fiebach, C. J., Schlesewsky, M., & Friederici, A. D. (2002). Separating syntactic memory costs and syntactic integration costs during parsing: The processing of German WH-questions. *Journal of Memory and Language, 47,* 250–272.

Friederici, A. D., Hahne, A., & Saddy, D. (2002). Distinct neurophysiological patterns reflecting aspects of syntactic complexity and syntactic repair. *Journal of Psycholinguistic Research, 31,* 45–63.

Friederici, A. D., Pfeifer, E., & Hahne, A. (1993). Event-related brain potentials during natural speech processing: Effects of semantic, morphological and syntactic violations. *Cognitive Brain Research, 1,* 183–192.

Friederici, A. D., Schlesewsky, M., & Fiebach, C. J. (2003). Wh-movement versus Scrambling: The brain makes a difference. In S. Karimi (Ed.), *Word order and scrambling* (pp. 325–344). Oxford: Blackwell Publishing.

Gibson, E. (1998). Linguistic complexity: Locality of syntactic dependencies. *Cognition, 68,* 1–76.

Gibson, E. (2000). Dependency locality theory: A distance-based theory of linguistic complexity. In A. Marantz, Y. Miyashita, & W. O'Neil (Eds.), *Image, language, brain* (pp. 95–126). Cambridge: MIT Press.

Greenhouse, S. W., & Geisser, S. (1959). On methods in the analysis of profile data. *Psychometrika, 24,* 95–112.

Grossman, M., Crino, P., Reivich, M., Stern, M. B., & Hurtig, H. I. (1992). Attention and sentence processing deficits in Parkinson's disease: The role of anterior cingulate cortex. *Cerebral Cortex, 2,* 513–525.

Gunter, T., Stowe, L. A., & Mulder, G. (1997). When syntax meets semantics. *Psychophysiology, 34,* 660–676.

Hagoort, P., Brown, C., & Groothusen, J. (1993). The syntactic positive shift (SPS) as an ERP measure of syntactic processing. *Language and Cognitive Processes, 8,* 439–483.

Hagoort, P., Brown, C., & Osterhout, L. (1999). The neurocognition of syntactic processing. In C. Brown & P. Hagoort (Eds.), *Neurocognition of language* (pp. 273–316). Oxford: Oxford University Press.

Jasper, H. H. (1958). The ten–twenty electrode system of the International Federation. *Electroencephalography and Clinical Neurophysiology, 10,* 371–375.

Kaan, E., Harris, A., Gibson, E., & Holcomb, P. (2000). The P600 as an index of syntactic integration difficulty. *Language and Cognitive Processes, 15,* 159–201.

Kaan, E., & Swaab, T. (2003). Repair, revision and complexity in syntactic analysis: An electrophysiological differentiation. *Journal of Cognitive Neuroscience, 15,* 98–110.

King, J. W., & Kutas, M. (1995). Who did what and when? Using word- and clause-level ERPs to monitor working memory usage in reading. *Journal of Cognitive Neuroscience, 1,* 378–395.

Kluender, R., & Kutas, M. (1993). Bridging the gap: Evidence from ERPs on the processing of unbounded dependencies. *Journal of Cognitive Neuroscience, 5,* 196–214.

Kutas, M., & Hillyard, S. A. (1983). Event-related brain potentials to grammatical errors and semantic anomalies. *Memory & Cognition, 11,* 539–550.

Lehmann, D., & Skrandies, W. (1984). Spatial analysis of evoked potentials in man: A review. *Progress in Neurobiology, 23,* 227–250.

Matzke, M., Mai, H., Nager, W., Russeler, J., & Münte, T. (2002). The costs of freedom: An ERP study of non-canonical sentences. *Clinical Neurophysiology, 113,* 844–852.

Münte, T. F., Heinze, H. J., & Mangun, G. R. (1993). Dissociation of brain activity related to syntactic and semantic aspects of language. *Journal of Cognitive Neuroscience, 5,* 335–344.

Neville, H. J., Nicol, J. L., Barss, A., Foster, K. I., & Garrett, M. F. (1991). Syntactically based sentence processing classes: Evidence from event-related brain potentials. *Journal of Cognitive Neuroscience, 3,* 151–165.

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

Osterhout, L., & Holcomb, P. J. (1992). Event-related brain potentials elicited by syntactic anomaly. *Journal of Memory and Language, 31,* 785–806.

Penke, M., Weyerts, H., Gross, M., Zander, E., Münte, T. F., & Clahsen, H. (1997). How the brain processes complex words: An event-related potential study of German verb inflections. *Cognitive Brain Research, 6,* 37–52.

Petrides, M., Alivisatos, B., Meyer, E., & Evans, A. C. (1993). Functional activation of the human frontal cortex during performance of verbal working memory tasks. *Proceedings of the National Academy of Sciences, U.S.A., 90,* 878–882.

Phillips, C., Kazanina, N., & Abada, S. H. (2005). ERP effects of the processing of syntactic long-distance dependencies. *Cognitive Brain Research, 22,* 407–428.

Rösler, F., Pechmann, T., Streb, J., Roder, B., & Henninghausen, E. (1998). Parsing of sentences in a language with varying word order: Word-by-word variations of processing demands are revealed by event-related brain potentials. *Journal of Memory and Language, 38,* 150–176.

Rösler, F., Puetz, P., Friederici, A. D., & Hahne, A. (1993). Event-related brain potentials while encountering semantic and syntactic constraint violations. *Journal of Cognitive Neuroscience, 5,* 345–362.

Rugg, M. D., & Coles, M. G. H. (1995). The ERP and cognitive psychology: Conceptual issues. In M. D. Rugg & M. G. H. Coles (Eds.), *Electrophysiology of mind* (pp. 27–39). Oxford: Oxford University Press.

Schlesewsky, M., Bornkessel, I., & Frisch, S. (2003). The neurophysiological basis of word order variations in German. *Brain and Language, 86,* 116–128.

Ueno, M., & Kluender, R. (2003). Event-related brain indices of Japanese scrambling. *Brain and Language, 86,* 243–271.