Interdisciplinary Information Sciences Vol. 22, No. 1 (2016) 123–141

ISSN 1340-9050 print/1347-6157 online
DOI 10.4036/iis.2016.R.04

Non-economical Verbal Information Processing Driven by a "Look-ahead" Strategy Under Poor Availability of Structural Information

Takahiro SOSHI[2,3,*] and Hiroko HAGIWARA[1,2,4]

[1]*Department of Language Sciences, Graduate School of Humanities, Tokyo Metropolitan University, Hachioji 192-0397, Japan*
[2]*Department of Humanity and Literatures, Faculty of Letters, Tokyo Metropolitan University, Hachioji 192-0397, Japan*
[3]*National Center of Neurology and Psychiatry, Kodaira 187-8553, Japan*
[4]*Research Center for Language, Brain and Genetics, Tokyo Metropolitan University, Hachioji 192-0397, Japan*

Natural language demonstrates discontinuous relationships between related words. Neurophysiology has reported that two types of brain electrical activities are related to discontinuous dependency. Sustained anterior negative potentials (SAN) are considered to be short-term memory storage for dislocated words. A late positivity (P600) appears concurrently with SAN decay, and is interpreted as the integration cost of a discontinuous word with a related word. Discontinuous verbal processing through SAN and P600 is considered to be unfavorable because of redundant neural resource consumption. However, SAN and P600 may reflect a global prediction-based strategy, which rather actively consumes memory resources to comprehend an overall meaning. We thus prepared four types of sentences (noun 1/adverb/noun 2/verb), which were modulated by grammatical information (case marker) and word order factors, and recorded brain potentials from Japanese participants performing a sentence comprehension task. Consistently with our prediction, P600 appeared and SAN disappeared upon the presentation of second nouns, but only in the canonical order including first subject words without a case marker. Hence, discontinuous verbal processing and its neural correlates should be re-considered in the context of interactions between local memory costs and global prediction strategies.

KEYWORDS: discontinuous structural dependency, global information processing, prediction strategy, sustained anterior negative cortical potential, late positive cortical potential

1. Introduction

A prominent feature of natural language is a discontinuity between constituent words [1], which is illustrated by the example, "*Who* did the student say that the teacher *admires*?" [2]. The question word "who" is located at the beginning of the sentence and is maximally separated from the related host "admires". The verb "admires" contains relational information for indispensable words, such as subjects and objects [3, 4]. We thus see that based on syntactic rules, when obligatory words do not appear adjacent to each other, they must occur in the same sentence, even if they are in "remote" or inverted positions [5]. Such discontinuous manipulations, or "displacements", are considered to be grounded in biologically determined natural language designs, such as interpretability of mental coding from other cognitive domains [6]. However, from the perspective of temporal verbal processing, a discontinuous or "filler-gap" dependency [7] creates a memory burden, which is needed to store and manipulate displaced words until the appearance of related words [8, 9]. Such high-loaded operations are contrary to another natural language design, called "cognitive economy", which is a constraint that leads to the avoidance of redundant neural resource consumption [10]. It has also been reported that displacement manipulations are difficult to recover in patients with neurological deficits, such as those with Broca's aphasia [11–14].

Design inconsistency between displacement manipulation and cognitive economy in natural language may be related to general cognitive function. Manipulation of dislocated words tends to be interpreted as local memory storage for ambiguous words (e.g., the unspecified subject or object status of "who" when processing the beginning of a sentence starting with the words "Who is ..."). However, local memory storage may be supported by a "look-ahead" global strategy [15–17]. A global strategy may violate local economy, because it leads to a high processing memory burden. However, it may be associated with the future goal of comprehending complex propositions under consideration [18].

Received March 17, 2016; Accepted September 1, 2016
*Corresponding author. E-mail: tksoshi@ybb.ne.jp

In fact, patients with Broca's aphasia often heavily depend on a prediction strategy (e.g., thematic hierarchy) when they cannot access explicit morpho-syntactical information [12, 19]. Since these patients tend to persist in a global strategy, they cannot correctly comprehend certain sentences, such as passive sentences with an inverted object-subject word order [11, 12, 14]. That is, a global verbal strategy may be clearly observable under an emergent condition: Verbal computation hence should be modeled not only as a bottom-up process from the perspective of local processing costs, but also as a top-down process, even if they are used to compensate for high memory resource consumption.

A computational integration model, the syntactic prediction locality theory (SPLT) or the dependency locality theory (DLT), proposes global verbal processing based on the notion of a memory resource [20, 21]. These look-ahead models introduce a numerical representation for prediction-based memory consumption and comprises two kinds of memory resource manipulation used for verbal comprehension. The first one is an "integration" component, which integrates input items into the current sentence structure, while the second is a "storage" component. The storage cost is represented by the "memory unit (MU)", which is determined by the number of "predicted" items needed to establish a minimal acceptable sentence. As shown in Table 1, the total MUs needed to reach the verb "admires" in the subject wh-question are 8 MUs, while the object wh-question consumes 11 MUs. When the initial "who" is encountered in both questions, the word required to construct a minimal sentence is an intransitive verb (e.g., "come"), which leads to the consumption of 1 MU. The difference between the two questions lies in the words following the complementizer "that". Since "the" and "teacher" appear after the word "that" in the object wh-question, a total of 3 MUs are required to predict the ensuing words, whereas 1 MU is used in the subject wh-question, where the word "admires" follows the word "that". That is, processing discontinuous dependency is considered to rely on local memory storage, which is driven by a global prediction strategy.

Table 1. Memory storage consumption for subject and object wh-questions in the prediction-based model.

Sentence type	Word order									
	1	2	3		4	5	6		7	
a. Subject wh-question	"*Who*	*did*	*the*	*student*	*say*	*that*	*Admires*		*the*	*teacher?*"
Storage cost (MU)	1	1	2	1	1	1	1		1	0
b. Object wh-question	"*Who*	*did*	*the*	*student*	*say*	*that*	*the*	*teacher*	*admires?*"	
Storage cost (MU)	1	1	2	1	1	1	2	1	0	

MU: memory unit.

Previous studies have utilized event-related potentials (ERPs) or neural electrical activity time-locked to an endogenous cognitive event [22] to investigate the neuropsychological correlates of storage and integration in discontinuous dependency [8, 9, 23–31]. Two surface brain electrical activity patterns have been frequently reported in this field. One is a sustained anterior negative potential (SAN), while the other is a late positive potential, which appears around 600 ms following the stimulus (P600).

SAN occurs after the appearance of a dislocated word and is believed to reflect the storage cost of a dislocated word in working memory [23]. The duration of SAN varies with different types of natural language. In a head-initial language, such as English (S–V–O order: subject/verb/object), SAN continues until a related verb is encountered. In a head-final language, such as German (S–O–V: subject/object/verb), SAN has been reported to terminate before the appearance of a verb [9, 31]. Therefore, SAN continues until the ambiguity of the nonintegrated word disappears. Functional interpretations of SAN as memory storage may be supported by: (1) the modulation of SAN by the distance between words in a discontinuous dependency [9, 30], or (2) the changes in the amplitude and the distribution of SAN in response to short-term memory spans [9].

The late positive potential, P600, on the other hand, appears after SAN decays, and is interpreted as a terminator of SAN. P600 is likely associated with the cost of integrating a dislocated word into the current representation in a discontinuous dependency [9, 25, 28, 31–33]. P600 can be observed independently of SAN [25], and has also been observed for both short- and long-separated dependencies [9, 30].

The neurophysiological substrates of discontinuous dependency, especially SAN, tend to be considered from the viewpoint of memory resource consumption. As indicated above, a discontinuous dependency may also lead to an attentional prediction related to an unintegrated word [8], as suggested by the attentional cueing model wherein attention and prediction interactively enhance neural activation for information processing [18, 34]. When we are exposed to canonical word orders with a high likelihood, we may form neural representations of an ordinary information flow, which may consume less neural resources. This has been observed in repetition suppression experiments for simple-stimulus priming [35]. In fact, a corpus-based study has reported that canonical sentences compose about 60% of all samples in a head-initial S–V–O language such as Finnish [36]. This suggests the existence of a prioritized neural representation for a canonical order [37]. Prediction of a canonical order may therefore be in accordance with an economy principle, such as the theory of minimal attachment, which is used to avoid complex verbal structure [38–40]. It may also follow a subject preference strategy, whereby an ambiguous first noun phrase (NP1) is assigned a subject role with structural priority [41, 42], which is in fact observed in neurological patients with Broca's aphasia [11, 12, 19]. However, various pragmatic concerns may require a change in ordinary information flow [43], such as agent (subject)/theme (object)/act (verb) in head-final S–O–V languages. These changes promote revised neural resource consumption for attentional prediction-based processing. Such alternations in information flow and related neural resource consumption may underlie the processing of discontinuous dependency [37].

To examine our hypothesis, we conducted a neurophysiological experiment. The Japanese language is characterized by a rich overt case system and allows a free word order. Specifically, case markers are attached immediately after nouns and provide structural and thematic information (e.g., an agent role for a subject) [44, 45]. Additionally, word orders freely change, as in the sentence "John-o/Mike-ga/home-ta", which literally translates to "John-object/Mike-subject/praise-past". Based on these properties of the Japanese language, we prepared four experimental conditions as a 2 (case and non-case) × 2 (canonical and non-canonical) factorial design (Fig. 1). The case-marked/canonical [$CAS_{(+)}/CAN_{(+)}$] condition is taken as the unmarked "baseline" condition with the least resource consumption. This condition includes sentences with a canonical S–O–V order and NP1 attached to a subject (Sbj) case marker ("-ga"). The canonical order is as follows: an NP1-Sbj precedes a second NP (NP2) with an object (Obj) case marker ("-o") [NP1-Sbj/Adverb (ADV)/NP2-Obj/verb phrase (VP)] (Fig. 1A). The case-marked/non-canonical [$CAS_{(+)}/CAN_{(-)}$] condition consists of non-canonical sentences with NP1s with an object case marker (NP1-Obj/ADV/NP2-Sbj/VP) (Fig. 1B). The non-case-marked/canonical [$CAS_{(-)}/CAN_{(+)}$] condition has a canonical word order and contains subject NP1s without a case marker (NP1/ADV/NP2-Obj/VP) (Fig. 1C). The non-case-marked/non-canonical [$CAS_{(-)}/CAN_{(-)}$] condition has a non-canonical word order with a non-case-marked object NP1 (NP1/ADV/NP2-Sbj/VP) (Fig. 1D).

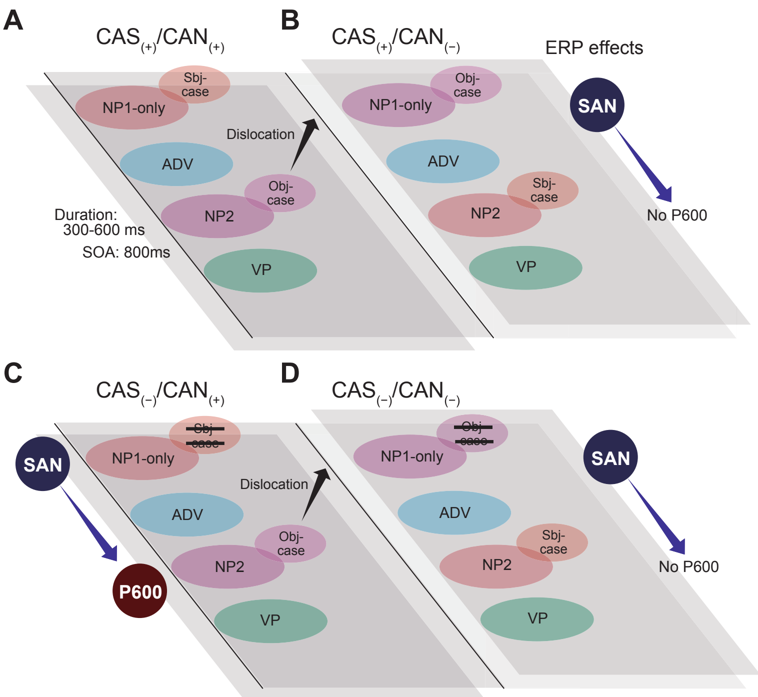

Fig. 1. The four experimental conditions. (A) The case-marked canonical condition [$CAS_{(+)}/CAN_{(+)}$] includes sentences with the first noun phrase (NP1) with a subject (Sbj) case marker and with a canonical word order [NP1-Sbj/adverb (ADV)/second NP with object case marker (NP2-Obj)/verb phrase (VP)]. (B) The case-marked non-canonical condition [$CAS_{(+)}/CAN_{(-)}$] consists of sentences with the NP1 with an object case marker and with a non-canonical word order (NP1-Obj/ADV/NP2-Sbj/VP). (C) The non-case-marked canonical condition [$CAS_{(-)}/CAN_{(+)}$] describes sentences with the non-case-marked subject NP1 and with a canonical word order (NP1/ADV/NP2-Obj/VP). (D) The non-case-marked non-canonical condition [$CAS_{(-)}/CAN_{(-)}$] describes sentences with the non-case-marked object NP1 and with a non-canonical word order (NP1/ADV/NP2-Sbj/VP). Stimulus durations varied from 300 to 600 ms, and the stimulus-onset-asynchrony was fixed to 800 ms. It was predicted that the three marked conditions [$CAS_{(+)}/CAN_{(-)}$, $CAS_{(-)}/CAN_{(+)}$, and $CAS_{(-)}/CAN_{(-)}$] would similarly elicit sustained anterior negativity (SAN) for memory storage of NP1s, but only the $CAS_{(-)}/CAN_{(+)}$ yielded P600 for prediction convergence of the object NP2.

The $CAS_{(+)}/CAN_{(+)}$ and $CAS_{(+)}/CAN_{(-)}$ conditions differ in their word order. Therefore, non-canonical $CAS_{(+)}/CAN_{(-)}$ sentences will yield an SAN for a non-canonical order, in contrast to $CAS_{(+)}/CAN_{(+)}$ sentences, which will not [31]. However, the presumed SAN may either reflect simple memory storage of dislocated words or it may reflect genuine attentional prediction-based processing. Our aim was to describe a condition that leads to genuine prediction-based processing, even in the absence of case information. We therefore added the following two conditions: $CAS_{(-)}$/$CAN_{(+)}$ and $CAS_{(-)}$/$CAN_{(-)}$. Because the NP1s do not possess case markers in either of the above conditions, they do not provide structural or thematic information by themselves and will not produce neural activities driven by case information. Based on the present 2×2 experimental design, $CAS_{(-)}$/$CAN_{(+)}$ sentences with non-case-marked NP1s may also yield SAN, in contrast to the $CAS_{(+)}$/$CAN_{(+)}$ sentences. However, in order to identify SAN as genuine prediction-based processing, it is important to clarify whether SANs disappear at certain hypothesized positions in $CAS_{(+)}$/$CAN_{(-)}$ and/or $CAS_{(-)}$/$CAN_{(+)}$ sentences. If P600s for the NP2s appear in both the $CAS_{(+)}$/$CAN_{(-)}$ condition (word order effect) and the $CAS_{(-)}$/$CAN_{(+)}$ condition (case effect) and similarly attenuate SAN, SAN may be interpreted to reflect memory storage of nonintegrated words (i.e., non-canonical objects and non-case-marked NP1s). However, if P600 is observed only in the $CAS_{(-)}$/$CAN_{(+)}$ condition, P600 should be interpreted to reflect convergence in a prediction strategy for a canonical word order. Thus SAN may also be associated with genuine prediction-based processing. To summarize, the elicitation of P600 in the $CAS_{(-)}$/$CAN_{(+)}$ and/or $CAS_{(+)}$/$CAN_{(-)}$ conditions may be crucial to clarifying the functional nature of SAN.

Our main hypothesis is represented in Fig. 1. SAN will appear in case-marked/non-canonical [$CAS_{(+)}$/$CAN_{(-)}$] and non-case-marked [$CAS_{(-)}$/$CAN_{(+)}$ and $CAS_{(-)}$/$CAN_{(-)}$] conditions as a prerequisite for the confirmation of structural representations at later computational stages (Fig. 1B-D). P600, on the other hand, is hypothesized to appear for the $CAS_{(-)}$/$CAN_{(+)}$ condition, which may drive genuine prediction-based processing in the absence of case information. SAN is hypothesized to disappear upon encountering the NP2 (Fig. 1C).

2. Methods

2.1 Participants

Twenty undergraduate students [14 women and 6 men; age (mean $\pm$ SD): 20.5 ± 1.6 years] participated in the experiment. All participants were native Japanese speakers and were right-handed (laterality quotient: 0.95 ± 0.1), as confirmed by the Edinburgh handedness inventory [46, 47]. Short-term verbal working memory required for the storage of words was assessed by the reading span test [48, 49], which was used to confirm whether participants could appropriately conduct sentence processing tasks. The participants had an average score of 3.5 ± 0.9, which indicates a normal verbal memory span. The participants self-reported that they had no history of brain injury due to accidents, no history of psychiatric treatment, and no current clinical medication. They all had normal or corrected-to-normal visual acuity. The participants provided informed consent in written form before the beginning of the experiment in accordance with the Declaration of Helsinki. The Human Subjects Ethics Committee of Tokyo Metropolitan University approved of all procedures. The participants received monetary compensation for their participation.

2.2 Experimental stimuli

We generated 40 sentences for each experimental condition [$CAS_{(+)}$/$CAN_{(+)}$, $CAS_{(+)}$/$CAN_{(-)}$, $CAS_{(-)}$/$CAN_{(+)}$, and $CAS_{(-)}$/$CAN_{(-)}$] (Table 2). All 160 sentences were simple sentences with a single clause and consisted of four words (NP1/ADV/NP2/VP). The NP1s were selected from 20 male and 20 female proper names with two kanji characters and three morae. Proper names for the NP1s were also used for the NP2s, so that individual sentences did not contain NP1s and NP2s with different gender properties and did not use the same name. The use of proper names most likely leads to semantic reversibility between the NP1 and the NP2 and deprives participants from a contextual guide for sentence comprehension. As observed in Table 2, combinations of the same names (e.g., "Mitsuo" and "Takeo") were not repeated to avoid a non-linguistic, goal-related strategy for the NP2 in repeated combinations. NP1s invariably contained the focus particle "-dake (only)". The reason for the use of this particle was that NPs with "-dake" may appear without a case marker, which enables us to investigate how case information affects verbal processing. Forty verbs were selected from a Japanese lexical database [50]. All verbs had high familiarity scores above 5.5 (5.92 ± 0.24) and included three characters with one kanji character and morae ranging from 3 to 5 (3.9 ± 0.55). To distract the participants' attention from the repeated use of similar sentence types, we produced 240 filler sentences including proper names with other focus particles "-shika (only)" or "-nomi (only)". Similar to "-dake", both particles indicate qualitative or quantitative limitation, and were expected to disrupt the participants' focus on "-dake". The fillers allowed object NPs to be attached to a subject or an object case marker (Table 2). A total of 400 sentences were divided into four lists (40 experimental and 60 filler sentences) and were randomly ordered. The order in which the four lists were presented was counterbalanced across the participants.

2.3 Procedures

Participants sat in an electrically shielded, soundproof room and faced a 17-inch cathode ray tube monitor, which was located 0.7 m away. Prior to the experimental trials, the participants performed a practice session to learn the procedure. The test trial began with a fixation symbol (∗∗∗∗) in black. The symbol was 5 cm wide and 1 cm tall and was displayed in the center of a light gray screen. All stimulus words were presented sequentially in black letters with

Table 2. Summary of experimental and filler sentences.

Conditions	Word order			
	1	2	3	4
	NP1	ADV	NP2	VP
$CAS_{(+)}/CAN_{(+)}$				
Japanese	"Mitsuo-dake-ga	kaigi-de	Takeo-o	seme-ta."
	Mitsuo-only-Sbj	conference-at	Takeo-Obj	accuse-Past
English	"Only Mitsuo	accused	Takeo	at the conference."
$CAS_{(+)}/CAN_{(-)}$				
Japanese	"Takeo-dake-o	kaigi-de	Hideo-ga	seme-ta."
	Takeo-only-Obj	conference-at	Hideo-Sbj	accuse-Past
English	"Hideo	accused	only Takeo	at the conference."
$CAS_{(-)}/CAN_{(+)}$				
Japanese	"Hideo-dake	kaigi-de	Yukio-o	seme-ta."
	Hideo-only-(Sbj)	conference-at	Yukio-Obj	Accuse-Past
English	"Only Hideo	accused	Yukio	at the conference."
$CAS_{(-)}/CAN_{(-)}$				
Japanese	"Yukio-dake	kaigi-de	Mitsuo-ga	seme-ta."
	Yukio-only-(Obj)	conference-at	Mitsuo-Sbj	accuse-Past
English	"Mitsuo	accused	only Yukio	at the conference."
Filler types	1	2	3	4
Filler 1				
Japanese	"Osamu-wa	bento-shika	gakko-ni	motteko-naka-tta."
	Osamu-Top	lunch box-only	school-to	bring-Neg-Past
English	"Osamu	brought	only a lunch box	to school."
Filler 2				
Japanese	"Shujii-nomi	Osamu-no	iukoto-ni	kanshinshi-ta"
	Family doctor-only	Osamu-Gen	utterance-by	be impressed with -past
English	"Only a family doctor	was impressed by	Osamu's utterance."	
Filler 3				
Japanese	"Kiyoshi-shika	Osamu-wa	itawara-naka-tta."	
	Kiyoshi-only'	Osamu-Top	care for-Neg-Past	
English	"Only Kiyoshi	cared for	Osamu."	
Filler 4				
Japanese	"Osamu-wa	Fusako-nomi	kangeishi-ta."	
	Osamu-Top	Fusako-only	greet-Past	
English	"Osamu	greeted	only Fusako."	
Filler 5				
Japanese	"Kenichi-wa	Shoichi-ga	shinsetsu-da-to	omo-tta."
	Kenichi-Top	Shoichi-Sbj	kind-Copula-Comp	consider-Past
English	"Kenichi'	considered	Shoichi	was kind."
Filler 6				
Japanese	"Kenichi-wa	Shoichi-o	shinsetsu-da-to	omo-tta."
	Kenichi-Top	Shoichi-Obj	kind-Copula-Comp	consider-Past
English	"Kenichi	considered	Shoichi	to be kind."

$CAS_{(+)}$ and $CAS_{(-)}$: case marked and non-marked; $CAN_{(+)}$ and $CAN_{(-)}$: canonical and non-canonical word order; NP1: the first noun phrase; ADV: adverb; NP2: the second noun phrase; VP: verb phrase; Sbj: subject case marker; Obj: object case marker; Top: topic case marker; Neg: negation (not); Gen: genitive case marker; Copula: copulative verb; Comp: clausal complementizer.

widths ranging from 4 cm to 6 cm and heights of 1.5 cm. The stimulus onset asynchrony (SOA) was 800 ms. Each stimulus word was presented for 300 to 600 ms. This duration was determined on the basis of a non-cumulative center-screen self-paced reading task performed about one month prior to the ERP experiment. The sentences in the reading task included 16 sentences containing four words each. Each word had three to six characters and morae evenly distributed in four positions in the sentence [e.g., "Hototogisu-ga (a little cuckoo-Sbj)/gikotinaku (unreadily)/yane-ni (on the roof)/toma-tta (perched)", which translates to "A little cuckoo unreadily perched on the roof"]. Using a linear regression analysis utilizing reading time data for individual words from 20 of the participants, we determined that it took on average about 100 ms (92 ms) to read one character with one mora. Therefore, a standard duration of 100 ms was used to calculate the stimulus durations of words in the ERP experiment.

At the end of the individual trials, a question sentence was presented and kept in the center of the screen until the participant responded using a "Yes" or "No" button. The question sentence briefly paraphrased a trial sentence in one of three manners [e.g., "Mitsuo-ga semeta?", translated to "Did Mitsuo blame (object person)?"; "Takeo-o semeta?", translated to "Did (subject person) blame Takeo?"; or "Kaigi-de semeta?", translated to "Did (subject person) blame (object person) at the conference?"]. The participants were instructed to judge whether the question sentence was consistent with the trial sentence. They were also instructed that concealed subjects, objects, and/or adverbs were assumed to be correctly memorized, and ignored them. Finally, they were required to press the buttons on the response pad as rapidly and accurately as possible.

2.4 Neurophysiological acquisition and analysis

Electroencephalograms (EEGs) were continuously recorded using 34 Ag/AgCl scalp electrodes, which were selected from 123 electrodes equidistantly mounted on the Quick-Cap 128 (Compumedics Neuroscan, Inc., Charlotte, NC) (Fig. 2). The central electrode (No. 17) was placed onto the Cz. The ground electrode was re-located to the midline anterior-frontal position (AFz) in order to maintain the Fz electrode location. Other electrodes, including thirteen symmetrically located electrodes in each hemisphere, were chosen at intervals of one line from the midline. One electrode from each line inside the 10/20 area was selected [51]. The distance between electrodes was about 5 cm. Three electrodes were placed at the outer canthus of each eye and below the left eye to obtain off-line vertical and horizontal electro-oculograms (EOGs) (vertical EOG: left upper canthus minus left lower canthus; horizontal EOG: left upper minus right upper canthus). All electrodes were referenced to the left mastoid on-line. The potential change in the right mastoid was recorded and used to re-reference to the linked mastoids off-line. The EEG was recorded with a sampling frequency of 250 Hz and was amplified using a band-pass filter ranging from 0 (DC) to 70 Hz. Impedances of all electrodes were set to below 5,000 Ω throughout the experimental trials.

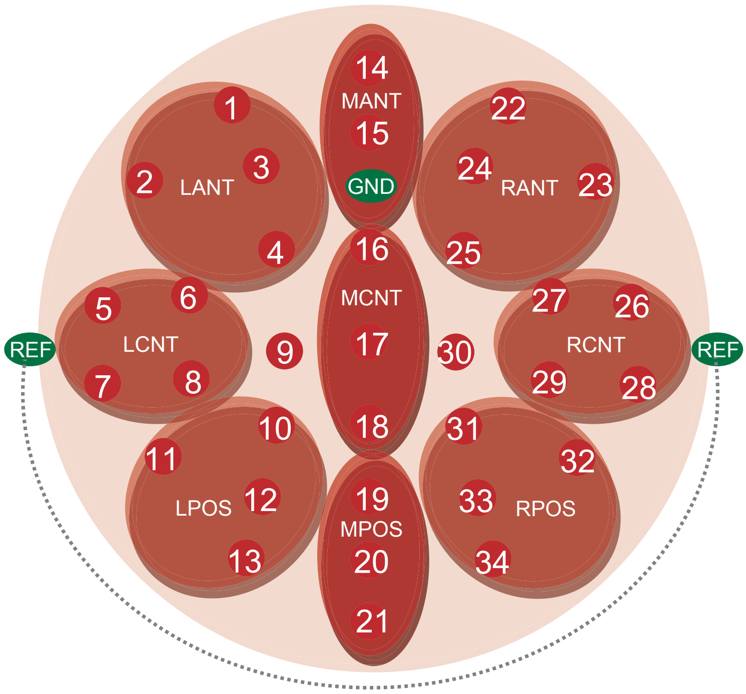

Fig. 2. Two-dimensional scalp map of 34 electrodes inside the 10/20 area. The clusters used for statistical analyses were circled (LANT: left anterior; LCNT: left central; LPOS: left posterior; RANT: right anterior; RCNT: right central; RPOS: right posterior; MANT: midline anterior; MCNT: midline central; MPOS: midline posterior). The ground (GND) electrode was located within the MANT area. Reference (REF) electrodes were placed at the bilateral mastoids, which were linked offline.

To examine SANs for dislocated and/or non-case-marked NP1s, a continuous EEG was filtered with band frequencies ranging from 0.3 to 40 Hz (FIR filter, −12 dB/octave, zero-phase shift). The present filter setting reduces three fourths of frequency powers, when lower and upper edge frequencies are halved or duplicated, respectively (0.15 and 80 Hz). The high-cut frequency of around 40 Hz has been used in previous studies [9, 29, 52]. The low-cut frequency of 0.3 Hz was applied to detrend immoderate baseline drifts, which tend to be observed in EEG recording with a DC amplifier [30]. Because SAN likely continues until the appearance of pre-verbal words (about 1600 ms duration in the present study) in head-final languages, the present low-cut frequency of 0.3 Hz may not strongly modify SAN with an about half-cycle duration. Continuous EEG was subsequently segmented into epochs starting from 200 ms prior to the presentation of the NP1 and ending at the NP2 for a multiword ERP (−200 to 2,400 ms). We also analyzed single word ERPs (−200 to 800 ms) for the NP2 in order to examine the convergence of a non-canonical and/or non-case-marked ambiguous NP1, and of the VP to examine the load for closing sentences. We utilized mean amplitudes in the 200 ms time interval before stimulus presentation to perform a baseline correction of the onset voltages of the EEG epochs. Epochs contaminated by EOGs and residual baseline drifts were eliminated using an automatic rejection process with a peak-to-peak amplitude of $\pm 75\,\mu V$. Epochs contaminated by residual artifacts with shapes obviously characteristic of vertical or horizontal EOGs were also removed manually. Epoch rejection rates were $23.8 \pm 1.4\%$ in the multiword analyses, while they were $14.2 \pm 5.8\%$ in the single word analyses. Grand average waveforms were smoothed using a moving average method (9 points $\times$ 4 ms $= 36$ ms) only to make visual inspection more convenient. To facilitate the inspection of surface potential distributions, two-dimensional scalp maps of difference potential values [$CAS_{(+)}/CAN_{(-)}$, $CAS_{(-)}/CAN_{(+)}$, and $CAS_{(-)}/CAN_{(-)}$ minus $CAS_{(+)}/CAN_{(+)}$] were constructed using a linear interpolation method (4 points) utilizing all 34 of the scalp electrodes. EEG analyses were performed using Scan 4.3 software (Compumedics Neuroscan, Inc., Charlotte, NC).

2.5 Statistical analysis

Behavioral data (response time and accuracy rate) were analyzed using a two-way repeated measures ANOVA with the within-participants factors of case [CAS: $CAS_{(+)}$ and $CAS_{(-)}$] and word order [WO: $CAN_{(+)}$ and $CAN_{(-)}$]. If interaction effects were significant, follow-up ANOVAs were performed for each factor.

To analyze neurophysiological effects, repeated measures ANOVAs utilizing non-smoothed raw amplitudes were performed for each 100 ms post-stimulus interval. Lateral and midline sites were tested separately. A grand ANOVA for the lateral site had the four within-subject factors of CAS, WO, hemisphere [HEM: left (LHEM) and right (RHEM)], and region of interest [ROI: anterior (ANT), central (CNT), and posterior (POS)]. To control for statistical power, each ROI contained four electrodes (Fig. 2). The overall ANOVA for the midline site included the three within-participants factors of CAS, WO, and ROI [midline anterior (MANT), midline central (MCNT), and midline posterior (MPOS)]. Since the MANT ROI contained the ground electrode, it had two electrodes, in contrast with the MCNT and MPOS ROIs, which had three electrodes. If we observed significant interaction effects between the CAS and/or WO factors in overall ANOVAs, we performed follow-up ANOVAs for each factor to elucidate simple CAS and/or WO effects. The Greenhouse-Geyser correction was not applied to correct for degree of freedom, as there were no violations of the sphericity assumption. To avoid confusion, we will report significant effects which finally yielded simple CAS and/or WO effects. Additionally, when significant effects appeared across multiple time windows, we will report results with maximal (max) statistical values. Statistical tests were performed using IBM SPSS Statistics software (IBM Corp., Tokyo, Japan).

3. Results

3.1 Behavioral results

Mean response accuracies (%) and response times (RTs) for correct responses for the four conditions used are presented in Table 3. The ANOVA for response accuracy indicated main CAS and WO effects [CAS: $F_{(1,19)} = 8.853$, $p < 0.01$; WO: $F_{(1,19)} = 13.516$, $p < 0.01$]. Averaged RTs showed a similar pattern. We observed significant main CAS and WO effects [CAS: $F_{(1,19)} = 6.526$, $p < 0.05$; WO: $F_{(1,19)} = 19.162$, $p < 0.001$]. As can be confirmed in Table 3, the non-case-marked [$CAS_{(-)}/CAN_{(+)}$ and $CAS_{(-)}/CAN_{(-)}$] and the non-canonical [$CAS_{(+)}/CAN_{(-)}$ and $CAS_{(-)}/CAN_{(-)}$] conditions were lower in response accuracy and longer in RT than the case-marked and the canonical conditions, respectively. These results indicate that sentences with non-case-marked and non-canonical NP1s are relatively difficult to comprehend.

Table 3. Means and standard deviations of response accuracies and response times (RT) in the four experimental conditions.

Conditions	Response accuracy (%)	RT (ms)
$CAS_{(+)}/CAN_{(+)}$	93 ± 5.7	1487 ± 437
$CAS_{(+)}/CAN_{(-)}$	80 ± 5.5	1716 ± 538
$CAS_{(-)}/CAN_{(+)}$	89 ± 5.5	1614 ± 473
$CAS_{(-)}/CAN_{(-)}$	75 ± 17.6	1870 ± 679

$CAS_{(+)}$ and $CAS_{(-)}$: case-marked and non-marked; $CAN_{(+)}$ and $CAN_{(-)}$: canonical and non-canonical word order.

3.2 Neurophysiological results

3.2.1 Multiword analysis

3.2.1.1 Sustained anterior negativity for the three consecutive phrases

A multiword analysis was conducted to examine the neurophysiological correlates of memory storage of nonintegrated (non-case-marked and non-canonical) NP1s, as in previous studies [8, 9, 23–28, 30, 31]. We found that SAN effects for both non-case-marked and non-canonical NP1s appeared from about 600 ms after the presentation of the NP1s, while SANs disappeared upon NP2 presentation only in the non-case-marked and canonical word order condition [CAS$_{(-)}$/CAN$_{(+)}$].

Multiword waveforms obtained in the four conditions between the NP1 and the NP2 were selected from nine lateral and midline ROIs and are superimposed and plotted in Fig. 3A. Potential maps of the difference amplitudes [CAS$_{(+)}$/CAN$_{(-)}$, CAS$_{(-)}$/CAN$_{(+)}$, and CAS$_{(-)}$/CAN$_{(-)}$ minus CAS$_{(+)}$/CAN$_{(+)}$] obtained 200 ms after the presentation of the NP1 are shown in Fig. 3B. The results of the grand ANOVAs are summarized in Table 4 for the NP1, in Table 5 for the ADV, and in Table 6 for the NP2. As seen in Figs. 3A and 3B, ERPs for the CAS$_{(+)}$/CAN$_{(-)}$, CAS$_{(-)}$/CAN$_{(+)}$, and

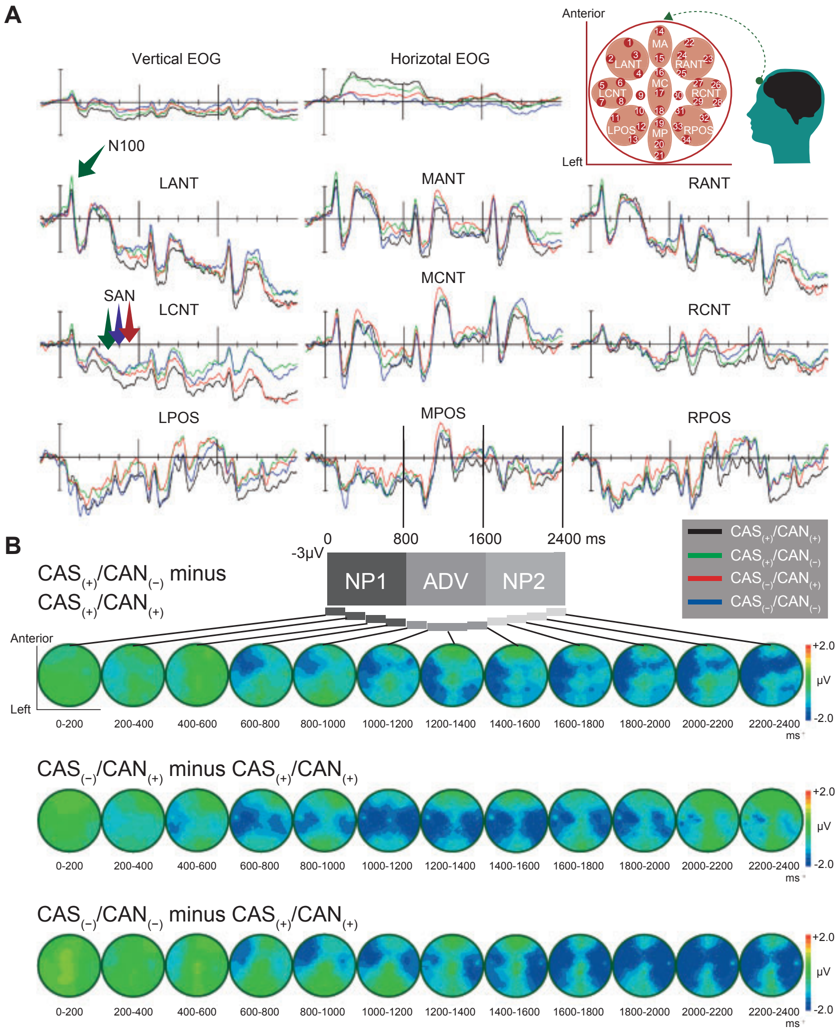

Fig. 3. Grand-averaged multiword event-related potential (ERP) results [the first noun phrase (NP1)/adverb (ADV)/the second noun phrase (NP2)] in the four conditions {case-marked/canonical [CAS$_{(+)}$/CAN$_{(+)}$; black line], case-marked/non-canonical [CAS$_{(+)}$/CAN$_{(-)}$; green line], non-case-marked/canonical [CAS$_{(-)}$/CAN$_{(+)}$; red line], and non-case-marked/non-canonical [CAS$_{(-)}$/CAN$_{(-)}$; blue line]}. (A) ERP waveforms from nine scalp regions (LANT: left anterior; LCNT: left central; LPOS: left posterior; RANT: right anterior; RCNT: right central; RPOS: right posterior; MANT: midline anterior; MCNT: midline central; MPOS: midline posterior) are presented for the four conditions. Sustained anterior negativity (SAN) effects for the CAS$_{(+)}$/CAN$_{(-)}$ (green), CAS$_{(-)}$/CAN$_{(+)}$ (red), and CAS$_{(-)}$/CAN$_{(-)}$ (blue), conditions compared to those in the CAS$_{(+)}$/CAN$_{(+)}$ (black) condition are illustrated. SAN effects are particularly apparent in the LANT and LCNT regions. Vertical and horizontal electro-oculograms are also plotted at the topmost positions. (B) Scalp potential maps of difference amplitudes [CAS$_{(+)}$/CAN$_{(-)}$, CAS$_{(-)}$/CAN$_{(+)}$, CAS$_{(-)}$/CAN$_{(-)}$ minus CAS$_{(+)}$/CAN$_{(+)}$] are represented for every 200 ms following NP1 presentation. More blue-colored areas indicate increased negative potential effects. Noticeably, the SAN gradually decreased during the NP2 epoch only for the CAS$_{(-)}$/CAN$_{(+)}$ condition.

Table 4. Summary of grand ANOVAs for the NP1 epoch in the multiword analysis.

Effect		Latency (ms)							
	df	0–100	100–200	200–300	300–400	400–500	500–600	600–700	700–800
Lateral									
CAS	1,19								
WO	1,19								
CAS × WO	1,19							6.702*	
CAS × HEM	1,19								
WO × HEM	1,19								
CAS × WO × HEM	1,19								
CAS × ROI	2,38		4.029*						
WO × ROI	2,38		4.757*						9.037**
CAS × WO × ROI	2,38				3.809*				
CAS × HEM × ROI	2,38								
WO × HEM × ROI	2,38								
CAS × WO × HEM × ROI	2,38								
Midline									
CAS	1,19		7.442*						
WO	1,19								
CAS × WO	1,19							5.892*	
CAS × ROI	2,38		3.963						
WO × ROI	2,38								4.183*
CAS × WO × ROI	2,38								

NP1: the first noun phrase; df: degree of freedom; CAS: case; WO: word order; HEM: hemisphere; ROI: region of interest; *: $p < 0.05$; **: $p < 0.01$.

Table 5. Summary of grand ANOVAs for the ADV epoch in the multiword analysis.

Effect		Latency (ms)							
	df	800–900	900–1000	1000–1100	1100–1200	1200–1300	1300–1400	1400–1500	1500–1600
Lateral									
CAS	1,19		6.123*	7.960*	6.249*	9.251**	6.443*	5.891*	
WO	1,19								
CAS × WO	1,19								
CAS × HEM	1,19								
WO × HEM	1,19								
CAS × WO × HEM	1,19								
CAS × ROI	2,38		3.951*	3.905*					
WO × ROI	2,38	4.321*	4.719*	6.467**					
CAS × WO × ROI	2,38								
CAS × HEM × ROI	2,38								
WO × HEM × ROI	2,38								
CAS × WO × HEM × ROI	2,38								
Midline									
CAS	1,19								
WO	1,19								
CAS × WO	1,19								
CAS × ROI	2,38								
WO × ROI	2,38								
CAS × WO × ROI	2,38				3.842*				

ADV: adverb; df: degree of freedom; CAS: case; WO: word order; HEM: hemisphere; ROI: region of interest; *: $p < 0.05$; **: $p < 0.01$.

$CAS_{(-)}/CAN_{(-)}$ conditions were characterized by SANs beginning around 600 ms after the presentation of the NP1, which is in contrast to the waveform obtained in the $CAS_{(+)}/CAN_{(+)}$ condition. SANs for the $CAS_{(+)}/CAN_{(-)}$ and the $CAS_{(-)}/CAN_{(-)}$ conditions continued until the end of the NP2 epoch, while the SAN for the $CAS_{(-)}/CAN_{(+)}$ gradually diminished during the NP2 interval.

During the NP1 interval (0–800 ms), significant effects of CAS and WO were first observed in the lateral ANT ROIs in the early time window from 100 to 200 ms [lateral: CAS × ROI, $F_{(2,38)} = 4.029$, $p < 0.05$; ANT: CAS,

Table 6. Summary of grand ANOVAs for the NP2 epoch in the multiword analysis.

Effect	df	Latency (ms) 1600–1700	1700–1800	1800–1900	1900–2000	2000–2100	2100–2200	2200–2300	2300–2400
Lateral									
CAS	1,19		4.893*	5.798*					
WO	1,19								4.602*
CAS × WO	1,19								
CAS × HEM	1,19								
WO × HEM	1,19								
CAS × WO × HEM	1,19								
CAS × ROI	2,38								
WO × ROI	2,38				4.018*				
CAS × WO × ROI	2,38								
CAS × HEM × ROI	2,38								
WO × HEM × ROI	2,38								
CAS × WO × HEM × ROI	2,38								
Midline									
CAS	1,19								
WO	1,19								
CAS × WO	1,19								
CAS × ROI	2,38								
WO × ROI	2,38								
CAS × WO × ROI	2,38								

NP2: the second noun phrase; df: degree of freedom; CAS: case; WO: word order; HEM: hemisphere; ROI: region of interest; *: $p < 0.05$.

$F_{(1,19)} = 7.554$, $p < 0.05$; WO × ROI: $F_{(2,38)} = 4.757$, $p < 0.05$; ANT: WO, $F_{(1,19)} = 5.333$, $p < 0.05$] (Table 4). Based on visually assessing the waveforms in Fig. 3A, we observe that N100 for the $CAS_{(+)}/CAN_{(-)}$ condition is more negative in anterior sites compared to the other conditions (green arrow in Fig. 3A). This likely contributes to the significant early effects in the lateral ANT ROI. The N100 effect for the $CAS_{(+)}/CAN_{(-)}$ condition is due to visual attention evoked by the non-canonical object NP1 which induces a local structural error, as similarly observed for a non-canonical initial word in German [52].

We observed a significant CAS × WO × ROI effect in the lateral ANOVA in the time window from 300 to 400 ms [$F_{(2,38)} = 3.809$, $p < 0.05$] (Table 4). Follow-up ANOVAs for each ROI revealed that there is a significant CAS effect in the canonical [$CAN_{(+)}$] condition, which appears in the lateral CNT ROI [CAS × WO: $F_{(1,19)} = 5.389$, $p < 0.05$; $CAN_{(+)}$: CAS, $F_{(1,19)} = 7.937$, $p < 0.05$] and the lateral POS ROI [CAS × WO: $F_{(1,19)} = 4.584$, $p < 0.05$; $CAN_{(+)}$: CAS, $F_{(1,19)} = 5.968$, $p < 0.05$]. These lateral effects are visually observed for the $CAS_{(-)}/CAN_{(+)}$ condition in Fig. 3B, which shows the appearance of lateralized centro-posterior negative effects for the non-case-marked NP1s.

Significant case-related effects were observed from 600 ms until the end of the NP1 stimulus in the lateral sites [600–700 ms: CAS × WO, $F_{(1,19)} = 6.702$, $p < 0.05$; $CAN_{(+)}$: CAS, $F_{(1,19)} = 10.405$, $p < 0.01$] and in the midline sites [600–700 ms: CAS × WO, $F_{(1,19)} = 5.892$, $p < 0.05$; $CAN_{(+)}$: CAS, $F_{(1,19)} = 5.215$, $p < 0.05$] (Table 4). The CAS effects are observed by comparing the $CAS_{(-)}/CAN_{(+)}$ and $CAS_{(+)}/CAN_{(+)}$ conditions in Fig. 3B. The WO effect also appeared in the lateral anterior ROI [700–800 ms: WO × ROI, $F_{(2,38)} = 9.037$, $p < 0.01$; ANT: WO, $F_{(1,19)} = 5.768$, $p < 0.05$], which is observed when comparing the $CAS_{(+)}/CAN_{(+)}$ and $CAS_{(+)}/CAN_{(-)}$ conditions in Fig. 3B.

During the presentation of the ADV (800–1600 ms), significant case-related effects were still observed in the lateral sites [CAS (max: 1200–1300 ms): $F_{(1,19)} = 9.251$, $p < 0.01$; CAS × ROI (max: 900–1000 ms): $F_{(2,38)} = 3.951$, $p < 0.05$] (Table 5). Follow-up analyses for each ROI indicated significant case effects in the lateral anterior and central ROIs [CAS: ANT, $F_{(1,19)} = 5.732$, $p < 0.05$; CNT: $F_{(1,19)} = 11.154$, $p < 0.01$]. These lateral negative effects in the non-case-marked conditions are illustrated in the comparison between the $CAS_{(+)}/CAN_{(+)}$ and $CAS_{(-)}/CAN_{(+)}$ conditions in Fig. 3B. Although potential maps of the difference ERPs between the $CAS_{(+)}/CAN_{(+)}$ and the $CAS_{(+)}/CAN_{(-)}$ conditions consistently showed negative effects for a non-canonical order, ANOVAs did not yield a significant WO effect after 1100 ms (Table 5). Since enhanced negative effects in the $CAS_{(-)}/CAN_{(+)}$ condition using ambiguous non-case-marked NP1s might attenuate word order-related effects, we also independently performed planned ANOVAs using the $CAS_{(+)}/CAN_{(+)}$ and $CAS_{(+)}/CAN_{(-)}$ conditions with the case-marked NP1s. We observed significant word order-related effects in the later intervals from 1100 to 1200 ms and from 1300 to 1400 ms in the lateral sites [WO (1300–1400 ms): $F_{(1,19)} = 4.709$, $p < 0.05$; WO × HEM (1100–1200 ms): $F_{(1,19)} = 4.645$, $p < 0.05$; LHEM: WO, $F_{(1,19)} = 16.078$, $p < 0.001$]. This demonstrates that left lateralized negative effects are indeed present in the $CAS_{(+)}/CAN_{(-)}$ condition, as illustrated in the comparison between $CAS_{(+)}/CAN_{(+)}$ and $CAS_{(+)}/CAN_{(-)}$ conditions in Fig. 3B.

During the NP2 interval (1600–2400 ms), the significant CAS effect disappeared until the 1900 ms time point [lateral (max: 1800–1900 ms): CAS, $F_{(1,19)} = 5.798$, $p < 0.05$] (Table 6). The SAN attenuation is illustrated by the comparison between the $CAS_{(-)}/CAN_{(+)}$ and $CAS_{(+)}/CAN_{(+)}$ conditions in Fig. 3B. In contrast, significant word order-related effects were observed after 1900 ms [lateral: WO × ROI (1900–2000 ms): $F_{(2,38)} = 4.018$, $p < 0.05$; ANT: WO, $F_{(1,19)} = 5.172$, $p < 0.05$; WO (2300–2400 ms), $F_{(1,19)} = 4.602$, $p < 0.05$]. To summarize, SANs for sentences with a canonical word order and a non-case-marked NP1 [$CAS_{(-)}/CAN_{(+)}$] terminated during early time windows, while SANs for sentences with a non-canonical word order [$CAS_{(+)}/CAN_{(-)}$ and $CAS_{(-)}/CAN_{(-)}$] did not disappear.

3.2.2 Single word analysis

We performed single-word analyses of the NP2 and VP epochs with a correction to the baseline immediately before their presentation in order to examine focused effects of computational congruency and computational load upon sentence closure or verb presentation. In summary, we found that P600 appeared only for the NP2 in the non-case-marked/canonical condition [$CAS_{(-)}/CAN_{(+)}$]. However, the non-canonical conditions without P600s upon NP2 presentation yielded negative effects for the final verb [$CAS_{(+)}/CAN_{(-)}$ and the $CAS_{(-)}/CAN_{(-)}$].

3.2.2.1 P600 for the second noun phrase

ERPs for the third NP2 in the four conditions are compared in Fig. 4A and potential maps of difference amplitudes for $CAS_{(+)}/CAN_{(-)}$, $CAS_{(-)}/CAN_{(+)}$, and $CAS_{(-)}/CAN_{(-)}$ minus $CAS_{(+)}/CAN_{(+)}$ are shown for every 100 ms in Fig. 4B. Statistical results are presented in Table 7.

The non-canonical word orders [$CAS_{(+)}/CAN_{(-)}$ and $CAS_{(-)}/CAN_{(-)}$] consistently showed a significant negative effect from early time points [lateral: WO (max: 700–800 ms), $F_{(1,19)} = 13.422$, $p < 0.01$; midline: WO (max: 0–100 ms), $F_{(1,19)} = 9.829$, $p < 0.01$]. The negative effect of the non-canonical order was distributed mainly in bilateral anterior-central ROIs and appeared throughout the entire epoch [lateral: WO × ROI (max: 400–500 ms), $F_{(2,38)} = 4.608$, $p < 0.05$; ANT: WO, $F_{(1,19)} = 18.903$, $p < 0.001$; CNT: WO, $F_{(1,19)} = 9.376$, $p < 0.01$].

Noticeable differences in case effects were observed in the middle time windows around 500 ms post-stimulus when comparing canonical and non-canonical orders. Specifically, we observed significant interaction effects between the CAS and the WO in the lateral sites [CAS × WO × HEM (max: 400–500 ms): $F_{(1,19)} = 5.372$, $p < 0.05$] and the midline sites [CAS × WO × ROI (400–500 ms): $F_{(2,38)} = 3.282$, $p < 0.05$]. Follow-up analyses indicated that when compared to the $CAS_{(+)}/CAN_{(+)}$, the $CAS_{(-)}/CAN_{(+)}$ condition yielded a case-related positive effect in the right hemisphere [CAS × WO (400–500 ms): $F_{(1,19)} = 14.459$, $p < 0.01$; $CAN_{(+)}$: CAS, $F_{(1,19)} = 9.015$, $p < 0.01$] and the MPOS ROI [CAS × WO (400–500 ms): $F_{(1,19)} = 10.131$, $p < 0.01$; $CAN_{(+)}$: CAS, $F_{(1,19)} = 5.246$, $p < 0.05$], shown by $CAS_{(-)}/CAN_{(+)}$ minus $CAS_{(+)}/CAN_{(+)}$ in Fig. 5B.

However, compared to the $CAS_{(+)}/CAN_{(-)}$ condition, the $CAS_{(-)}/CAN_{(-)}$ condition yielded a case-related negative potential effect in the right hemisphere [CAS × WO (400–500 ms): $F_{(1,19)} = 14.459$, $p < 0.001$; $CAN_{(-)}$: CAS, $F_{(1,19)} = 7.507$, $p < 0.05$] and in the midline centro-posterior ROI [CAS × WO × ROI (400–500 ms): $F_{(2,38)} = 3.282$, $p < 0.05$; MCNT: CAS × WO, $F_{(1,19)} = 10.660$, $p < 0.01$; $CAN_{(-)}$: CAS, $F_{(1,19)} = 12.086$, $p < 0.01$; MPOS: CAS × WO, $F_{(1,19)} = 10.131$, $p < 0.01$; $CAN_{(-)}$: CAS, $F_{(1,19)} = 4.933$, $p < 0.05$].

To summarize, we observed a right posterior P600 only in sentences with a non-case-marked, canonical subject NP1 [$CAS_{(-)}/CAN_{(+)}$]. The non-canonical conditions generally continued to produce anterior-central negative effects. Additionally, when the non-case-marked condition had a non-canonical order [$CAS_{(-)}/CAN_{(-)}$], it increased posterior-dominant negative effects.

3.2.2.2 Negative effects for the final verb

Averaged waveforms for the final verbs are presented in Fig. 5A. Potential maps of difference amplitudes [$CAS_{(+)}/CAN_{(-)}$, $CAS_{(-)}/CAN_{(+)}$, $CAS_{(-)}/CAN_{(-)}$ minus $CAS_{(+)}/CAN_{(+)}$] for every 100 ms interval are presented in Fig. 5B. The grand ANOVA results are summarized in Table 8. To summarize the findings, in non-canonical orders [$CAS_{(+)}/CAN_{(-)}$ and $CAS_{(-)}/CAN_{(-)}$], the verb generally yielded negative effects from around 300 ms post-stimulus, which was in contrast to what we observed with canonical verbs [$CAS_{(+)}/CAN_{(+)}$ and $CAS_{(-)}/CAN_{(+)}$]. Additionally, in contrast to the case-marked condition [$CAS_{(+)}/CAN_{(-)}$], verbs in the non-case-marked condition [$CAS_{(-)}/CAN_{(-)}$] enhanced negative effects, but only when the sentences were ordered non-canonically.

In overall ANOVAs, effects related to a non-canonical word order [$CAS_{(+)}/CAN_{(-)}$ and $CAS_{(-)}/CAN_{(-)}$] began to appear in bilateral posterior sites around 300 ms post-stimulus and continued until the end of the epoch in lateral [WO × ROI (300–400 ms): $F_{(2,38)} = 5.504$, $p < 0.05$; WO (500–600 ms): $F_{(1,19)} = 4.925$, $p < 0.05$] and midline sites [WO × ROI (max: 700–800 ms): $F_{(2,38)} = 7.590$, $p < 0.01$]. This word order-related negative effect was mainly localized to posterior sites, as shown by the follow-up ANOVA for the midline site [WO: MPOS, $F_{(1,19)} = 8.402$, $p < 0.01$]. The posterior-dominant negative effects for the $CAS_{(+)}/CAN_{(-)}$ and $CAS_{(-)}/CAN_{(-)}$ conditions are reproduced in Fig. 5B.

Because the highest lateral interaction (CAS × WO × HEM × ROI) was observed from 600 to 700 ms [$F_{(2,38)} = 3.348$, $p < 0.05$], we also analyzed case-related effects. A significant CAS × WO effect appeared in the right ANT and CNT ROIs [RHEM: CAS × WO × ROI, $F_{(2,38)} = 4.845$, $p < 0.05$; ANT: CAS × WO, $F_{(1,19)} = 10.686$,

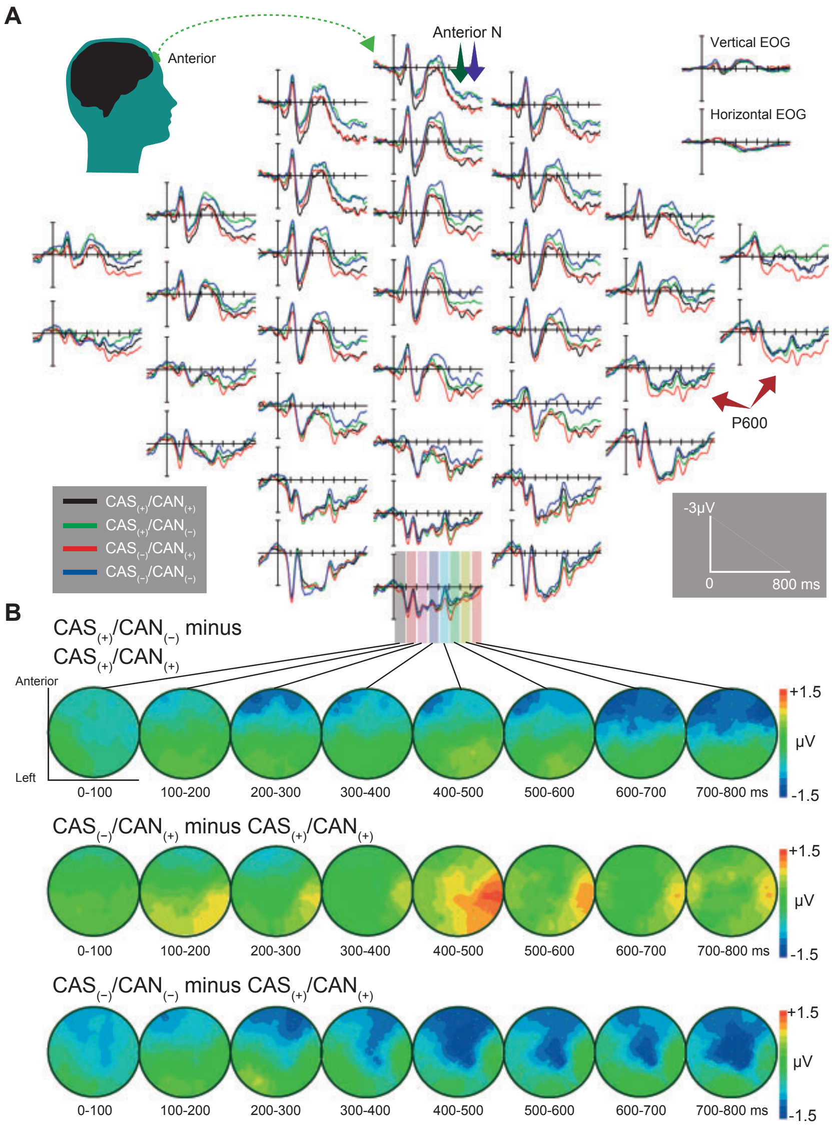

Fig. 4. Grand-averaged single word event-related potential (ERP) results of the second noun phrase (NP2) in the four conditions {case-marked/canonical [$CAS_{(+)}/CAN_{(+)}$; black line], case-marked/non-canonical [$CAS_{(+)}/CAN_{(-)}$; green line], non-case-marked/canonical [$CAS_{(-)}/CAN_{(+)}$; red line], and non-case-marked/non-canonical [$CAS_{(-)}/CAN_{(-)}$; blue line]}. (A) ERP waveforms from nine regions of interest (ROI: LANT, left anterior; LCNT: left central; LPOS: left posterior; RANT: right anterior; RCNT: right central; RPOS: right posterior; MANT: midline anterior; MCNT: midline central; MPOS: midline posterior) are plotted for each of the four conditions. A P600 specific for the $CAS_{(-)}/CAN_{(+)}$ (red) is clearly observed in right posterior-dominant sites. Anterior negativity (N) effects for the $CAS_{(+)}/CAN_{(-)}$ (green) and $CAS_{(-)}/CAN_{(-)}$ (blue) conditions, which are different from those of the $CAS_{(+)}/CAN_{(+)}$ (black) condition, are also illustrated. Vertical and horizontal electro-oculograms are also plotted at the right topmost positions. (B) Scalp potential maps of difference amplitudes [$CAS_{(+)}/CAN_{(-)}$, $CAS_{(-)}/CAN_{(+)}$, $CAS_{(-)}/CAN_{(-)}$ minus $CAS_{(+)}/CAN_{(+)}$] are represented for every 100 ms interval after the presentation of the NP2. Blue- and red-colored areas indicate increased negative and positive potential effects, respectively. The P600 for prediction convergence appears only for the $CAS_{(-)}/CAN_{(+)}$ condition.

$p < 0.01$; CNT: CAS × WO, $F_{(1,19)} = 4.425$, $p < 0.05$]. We also conducted subsequent ANOVAs for a simple CAS effect in the two ROIs. A significant CAS effect was observed when comparing the $CAS_{(-)}/CAN_{(-)}$ and $CAS_{(+)}/CAN_{(-)}$ conditions in the right ANT ROI [CAS × WO: $F_{(1,19)} = 10.686$, $p < 0.01$; $CAN_{(-)}$: CAS, $F_{(1,19)} = 5.966$, $p < 0.05$]. This can also be visually confirmed in Fig. 5B, wherein the negative effects for $CAS_{(-)}/CAN_{(-)}$ condition are more anteriorly distributed than those for the $CAS_{(+)}/CAN_{(-)}$ condition.

4. Discussion

4.1 Overview of results

Displacement manipulation alters the ordinary information flow seen in a canonical word order and consumes greater-than-normal neural resources for the communication of the intended meaning. Thus, inconsistencies between the displacement architecture and the principle of economy in natural language underlie discontinuous verbal

Table 7. Summary of grand ANOVAs of the NP2 in the single-word analysis.

Effect		Latency (ms)							
	df	0–100	100–200	200–300	300–400	400–500	500–600	600–700	700–800
Lateral									
CAS	1,19								
WO	1,19	11.415**	6.747*	5.683*	6.120*	12.493**	12.789**	10.035**	13.422**
CAS × WO	1,19					11.400**	4.798*		
CAS × HEM	1,19								
WO × HEM	1,19								
CAS × WO × HEM	1,19			4.732*		5.372*			
CAS × ROI	2,38		5.273*						
WO × ROI	2,38			4.392*		4.608*			
CAS × WO × ROI	2,38								
CAS × HEM × ROI	2,38								
WO × HEM × ROI	2,38								
CAS × WO × HEM × ROI	2,38								
Midline									
CAS	1,19								
WO	1,19	9.829**				8.508**	7.559*	5.528*	7.432*
CAS × WO	1,19					9.056**			
CAS × ROI	2,38								
WO × ROI	2,38								
CAS × WO × ROI	2,38								

NP2: the second noun phrase; df: degree of freedom; CAS: case; WO: word order; HEM: hemisphere; ROI: region of interest; *: $p < 0.05$; **: $p < 0.01$.

dependency. Here we hypothesized that such design inconsistencies were related to the requirements of a global prediction strategy. To test this hypothesis, we performed a neurophysiological experiment to clarify the precise functional natures of neural correlates of discontinuous dependency.

In the context of discontinuous dependency, it has heretofore frequently been argued that SAN is associated with the memory storage cost of dislocated words, and that P600 is related to the costs of integrating a dislocated word into the current structure. By manipulating the case and word order factors, we aimed to examine whether SAN was simply related to the storage cost of a dislocated word and whether P600 reflected integration costs.

Our behavioral experiments confirmed that sentences with a non-case-marked initial word and a non-canonical word order were difficult to comprehend and required longer processing times. These results support a cognitive non-economy in discontinuous verbal processing.

Our neurophysiological results indicated the presence of SANs for case and word order factors, but revealed differences between the two factors. The SAN disappeared upon the presentation of the third NP2, but only for sentences with a canonical word order and non-case-marked NP1s [$CAS_{(-)}/CAN_{(+)}$]. We observed a P600 upon the presentation of the third NP2, suggesting that the SAN terminated at the same time that the P600 appeared. On the other hand, while the non-canonical conditions [$CAS_{(+)}/CAN_{(-)}$ and $CAS_{(-)}/CAN_{(-)}$] continued to increase anterior-central negative potentials upon the presentation of the NP2, they did not produce a P600. Additionally, in contrast to the $CAS_{(-)}/CAN_{(+)}$ condition, the non-canonical conditions elicited posterior-dominant negativities upon the presentation of the final verbs. Thus, when a canonical order sentence begins with a non-case-marked NP1, it affects neural verbal processing differently when compared to other marked sentences.

4.2 Sustained anterior negativity as genuine prediction-based verbal processing

The case and word order factors had different sustained negative effects beginning from 600 ms after the presentation of the NP1. The non-case-marked sentences elicited effects that appeared consistently throughout the ADV interval. However, we could observe the word order effect when we used a planned comparison between the two case-marked conditions with different word orders [$CAS_{(+)}/CAN_{(+)}$ and the $CAS_{(+)}/CAN_{(-)}$]. This indicates that when brain activation states are included in overall comparisons of non-case-marked and canonical conditions [$CAS_{(-)}/CAN_{(+)}$], it may be difficult to detect a SAN for sentences with a non-canonical order. This may be due to the fact that the $CAS_{(-)}/CAN_{(+)}$ condition may also elicit a SAN, which would conceal a word order effect. The SAN was absent until 300 ms post-stimulus during the third NP2 epoch, but only in the non-case marked, canonical condition, which supports our prediction.

Our results seem to be inconsistent with previous studies, which have reported SANs for discontinuous dependencies in head-final languages (German: [9]; Japanese: [31]). The pre-verb subject nouns in the present $CAS_{(+)}/CAN_{(-)}$

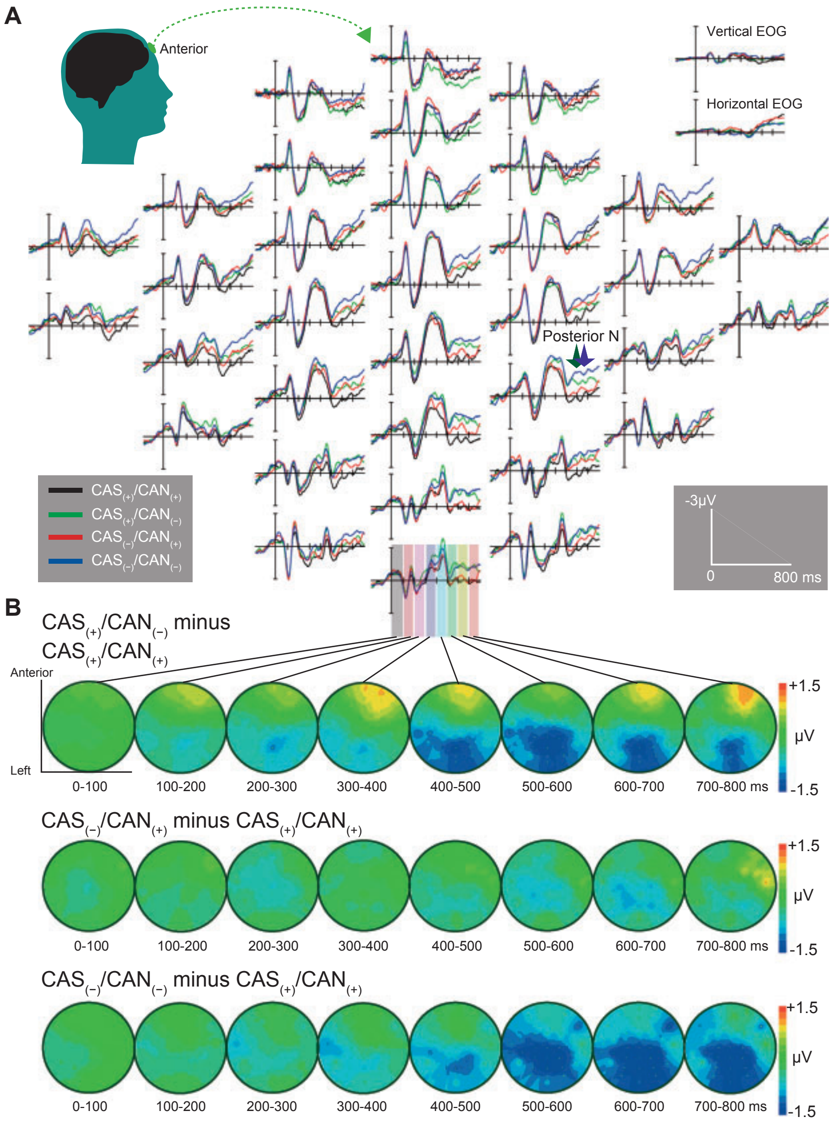

Fig. 5. Grand-averaged single word event-related potential (ERP) results of the final verb phrase (VP) in the four conditions {case-marked/canonical [CAS$_{(+)}$/CAN$_{(+)}$; black line], case-marked/non-canonical [CAS$_{(+)}$/CAN$_{(-)}$; green line], non-case-marked/canonical [CAS$_{(-)}$/CAN$_{(+)}$; red line], and non-case-marked/non-canonical [CAS$_{(-)}$/CAN$_{(-)}$; blue line]}. (A) ERP waveforms from the nine regions of interest (ROI: LANT, left anterior; LCNT: left central; LPOS: left posterior; RANT: right anterior; RCNT: right central; RPOS: right posterior; MANT: midline anterior; MCNT: midline central; MPOS: midline posterior) are plotted for each of the four conditions. Posterior negativity (N) effects for the non-canonical CAS$_{(+)}$/CAN$_{(-)}$ (green) and CAS$_{(-)}$/CAN$_{(-)}$ (blue) conditions are clearly observed in posterior-dominant areas. Vertical and horizontal electro-oculograms are also plotted at the right topmost positions. (B) Scalp potential maps of difference amplitudes [CAS$_{(+)}$/CAN$_{(-)}$, CAS$_{(-)}$/CAN$_{(+)}$, CAS$_{(-)}$/CAN$_{(-)}$ minus CAS$_{(+)}$/CAN$_{(+)}$] are shown for every 100 ms interval after the presentation of the VP. Dark blue-colored areas indicate increased negative potential effects. Enhanced negative effects of the semantic integration cost for the final verb do not appear in the CAS$_{(-)}$/CAN$_{(+)}$ condition.

condition, although similar to the pre-verb subject nouns in the previous long-distance dependency condition [31], did not lead to SAN termination. However, in the non-case marked, canonical condition [CAS$_{(-)}$/CAN$_{(+)}$], the object NP2 led to an attenuated SAN. That is, the SAN can disappear in both non-canonical and canonical word orders.

This conflicting data in the present work and previous studies suggests that SAN may not only represent the memory storage cost of a dislocated word. Instead, the SAN should be interpreted as a shared functional property of dislocated, case-marked NPs and non-dislocated, non-case-marked NPs. These two types of NP1 are not integrated into current structural representations when they first appear. The non-case-marked NP1 in CAS$_{(-)}$/CAN$_{(+)}$ sentences cannot be integrated into the current verbal representation, as there is no case marker to provide structural information. Similarly, the dislocated, case-marked object NPs in the previous study are also not readily integrated, as overall structural information only becomes clear when the object NPs are encountered. Such non-integrated properties of the NP1s described in the present work and in previous studies may yield a SAN, irrespective of differences in the dislocated properties of the NP1s.

Table 8. Summary of grand ANOVAs of the VP in the single-word analysis.

Effect		Latency (ms)							
	df	0–100	100–200	200–300	300–400	400–500	500–600	600–700	700–800
Lateral									
CAS	1,19								
WO	1,19						4.925*		
CAS × WO	1,19								
CAS × HEM	1,19								
WO × HEM	1,19								
CAS × WO × HEM	1,19								
CAS × ROI	2,38								
WO × ROI	2,38				5.504*				
CAS × WO × ROI	2,38					4.288*			
CAS × HEM × ROI	2,38								
WO × HEM × ROI	2,38								
CAS × WO × HEM × ROI	2,38							3.348*	
Midline									
CAS	1,19								
WO	1,19								
CAS × WO	1,19								
CAS × ROI	2,38								
WO × ROI	2,38				6.192*	5.389*	5.749*	7.517**	7.590**
CAS × WO × ROI	2,38								

VP: verb phrase; df: degree of freedom; CAS: case; WO: word order; HEM: hemisphere; ROI: region of interest; *: $p < 0.05$; **: $p < 0.01$.

A more crucial finding regarding the functional nature of SAN is that it may have different timespans in the CAS$_{(-)}$/CAN$_{(+)}$, CAS$_{(+)}$/CAN$_{(-)}$, and CAS$_{(-)}$/CAN$_{(-)}$ conditions. The sustained negative effect in the CAS$_{(-)}$/CAN$_{(+)}$ condition disappeared at an early time point when the NP2 was presented, while SANs in the CAS$_{(+)}$/CAN$_{(-)}$ and the CAS$_{(-)}$/CAN$_{(+)}$ conditions were still observed during the NP2 interval. This result is inconsistent with those in previous studies using head-final languages. The SAN for remotely dislocated object wh-words disappeared upon presentation of the pre-verb subject NP in German [9]. The SAN for a long-dislocated object NP also remained active until the pre-verb subject NP appeared [31]. Thus, non-canonical word orders led to SANs that ended at pre-verb positions in previous studies using head-final languages. These findings suggest that the duration of the SAN is not simply determined by the dislocation property of words. If a subject preference or a case filling strategy is applied to a non-case-marked ambiguous word in a canonical order [42], an object NP2 appears at the pre-verb position as expected, which in turn satisfies the prediction-based strategy in the CAS$_{(-)}$/CAN$_{(+)}$ condition. Based on the notion that a preference strategy is coupled with the prediction of disambiguating words, such as object words, a SAN may be strongly representative of the prediction of disambiguating targets, as argued in the DLT [21].

We can easily explain the continuously-observed SAN for the NP2 in the CAS$_{(-)}$/CAN$_{(-)}$ condition based on a prediction-based account. The NP1 in the CAS$_{(-)}$/CAN$_{(-)}$ condition does not have a case marker. Thus, a subject preference strategy may be at work in the CAS$_{(-)}$/CAN$_{(-)}$ condition. The NP2 in the CAS$_{(-)}$/CAN$_{(-)}$ condition, however, unexpectedly has a subject case marker. The processing of the non-canonical relation between the NP1 and the unexpected NP2 would take place following the failure of the initial prediction, which would induce a higher working memory load [8]. In consequence, the SAN may not terminate at the NP2 in the CAS$_{(-)}$/CAN$_{(-)}$ condition.

Non-decay of the SAN in the CAS$_{(+)}$/CAN$_{(-)}$ condition is apparently more difficult to explain. Since the NP1s have an overt object case marker, the following subject NP2 is easily predicted, especially since it is expected to appear in an experimental context. Therefore, the violation of prediction or subject preference would not be a main reason for the continuity of the SAN in the CAS$_{(+)}$/CAN$_{(-)}$ condition. Most likely, the participants have already established default neural representations for a canonical word order with a high likelihood during the course of their lives. Therefore, non-canonical word orders may conflict with default neural representations and lead to high processing loads, even when case information is present. Thus, the persistence of the SAN in the CAS$_{(+)}$/CAN$_{(-)}$ condition may be consistent with the idea of an increased working memory load due to a dislocated word order [8].

4.3 P600: integration cost versus prediction congruity

Single word analysis of the NP2 clearly demonstrates differences in ERP effects among the CAS$_{(+)}$/CAN$_{(-)}$, the CAS$_{(-)}$/CAN$_{(+)}$ and the CAS$_{(-)}$/CAN$_{(-)}$ conditions. In contrast to the NP2 in the CAS$_{(+)}$/CAN$_{(+)}$ condition, the NP2 in the CAS$_{(-)}$/CAN$_{(+)}$ condition only yields a positive effect. This positive effect likely corresponds to P600. However, the NP2s in the CAS$_{(+)}$/CAN$_{(-)}$ and CAS$_{(-)}$/CAN$_{(-)}$ conditions increased negative potentials, which were mainly distributed in anterior-central sites and began at early time points. Additionally, in contrast to the NP2 in the $CAS_{(+)}/CAN_{(-)}$ condition, the NP2 in the $CAS_{(-)}/CAN_{(-)}$ condition enhanced negative effects in the right lateral and midline central-posterior sites around 400 ms post-stimulus.

SAN is reported to be attenuated by the elicitation of P600 [9, 30, 31]. Therefore, the P600 observed in the present study also likely erased the SAN in the $CAS_{(-)}/CAN_{(+)}$ condition. Nevertheless, there is a noticeable difference in P600 between the present work and previous studies. P600 has generally been observed for complex sentences with discontinuous dependencies and is associated with the dislocation of words in particular [9, 25, 30, 31, 53]. However, P600 was only observed in $CAS_{(-)}/CAN_{(+)}$ sentences, which had a canonical word order without inversion of a word order. The reason for the increased P600 in the $CAS_{(-)}/CAN_{(+)}$ condition should therefore be clarified.

Based on the prediction-based account of SAN, it may be appropriate to interpret the P600 in the present study as a sign of the congruent prediction of an object NP, rather than as an integration cost [25]. That is, when predicted types of words appear, they disambiguate the roles of non-integrated words and may lead to the production of a P600. This argument is also supported by our findings regarding the negative effects observed in the $CAS_{(+)}/CAN_{(-)}$ and $CAS_{(-)}/CAN_{(-)}$ conditions. Although the negative effects for both of the conditions are similarly distributed in anterior sites, the negative effects were enhanced in the $CAS_{(-)}/CAN_{(-)}$ condition when compared to the $CAS_{(+)}/CAN_{(-)}$ condition. The NP2 with a subject case marker in the $CAS_{(-)}/CAN_{(-)}$ condition, which is different from the NP2 in the $CAS_{(+)}/CAN_{(-)}$ condition, is inconsistent with a subject preference and prediction strategy, whereby a disambiguating NP2 is prospectively assigned an object role. This prediction violation may thus lead to additional negative effects in the right lateral and midline posterior regions. To summarize, the NP2s in the $CAS_{(+)}/CAN_{(-)}$ and $CAS_{(-)}/CAN_{(-)}$ conditions may commonly induce a working memory load due to their integration difficulty in non-canonical word orders [8], which would result in not P600, but rather anterior-dominant negativity. Additionally, the violation of a subject preference strategy likely yielded the additional posterior dominant negative effects observed in the non-canonical, non-case-marked $CAS_{(-)}/CAN_{(-)}$ condition.

The P600 in the present study had a right posterior-dominant scalp distribution, which is different from the left fronto-central dominant P600 observed in the previous study [31]. The frontal-dominant P600 may be mainly related to structural integration [31], which distinguishes it from the posterior P600 [30]. Thus the kinds of verbal processes eliciting the right posterior-dominant P600 in the present study should be further clarified.

The posterior P600 associated with integration has frequently been reported in previous studies of western languages [9, 25, 30]. The present posterior P600, therefore, should be interpreted in accordance with previous studies. Although it has been widely accepted that the posterior P600 is related to linguistic integration manipulation [9, 25, 31], there are still questions regarding the nature of the integration processes reflected by the posterior P600. Kaan *et al.* [25] defined the notion of syntactic integration based on the DLT as a process through which input words are combined according to a predicted structure. Fiebach *et al.* [9] similarly suggest that the structural integration of a dislocated word occurs through the use of case information. Phillips *et al.* narrowly interpret the integration P600 as the process of the integration of thematic roles. On the other hand, the present P600, which was observed only in the $CAS_{(-)}/CAN_{(+)}$ condition, strongly supports the idea of a prediction-based P600 [25]. Here, we hypothesize that integration operations associated with P600s should be divided into sub-processes, as suggested by Phillips *et al.* (2005) [30]. A crucial difference between the present work and previous studies is its use of case marking [31]. The NP1 in the $CAS_{(-)}/CAN_{(+)}$ condition does not have a case marker, while the non-canonical NP1 in the previous study had an object case marker. That is, the processing of the NP1 in the $CAS_{(-)}/CAN_{(+)}$ condition was not driven by case information, but may be strongly affected by an attentional prediction of a disambiguating object NPs. However, the non-canonical object NP1 in the previous study already had a case marker when it was presented. Therefore, the integration process in the previous study was less dependent on a prediction strategy, when compared to the present study. Taken together, our data indicate that in contrast to the right posterior P600, the left fronto-central P600 may reflect the structural integration cost of congruent words in the context of non-canonical verbal processing.

To summarize overall, the P600 in the $CAS_{(-)}/CAN_{(+)}$ condition may reflect prediction congruity. On the other hand, the overlapping anterior negative effects for the $CAS_{(+)}/CAN_{(-)}$ and $CAS_{(-)}/CAN_{(-)}$ conditions may rather be related to a genuine integration cost of a non-canonical word order, which increases working memory costs.

4.4 Negative effects for final words in a non-canonical word order

The non-canonical conditions [$CAS_{(+)}/CAN_{(-)}$ and $CAS_{(-)}/CAN_{(-)}$] yielded posterior-dominant negative effects about 300 ms post-stimulus. The distribution of the negative effect differs from the anterior focused negative effects of the final verb in a non-canonical sentence [28], but is similar to the distributional effects of the final verb in a long-distance non-canonical condition reported in a previous study [31]. There are two possible interpretations for the present posterior negative effects. The first plausible interpretation is that the negative effect is due to a contingent negative variation or an expectancy wave (CNV: see for review, [54]) despite having a posterior dominant distribution. After the disappearance of the final verb, a CNV may increase until a task-related question is presented for an individual trial [24]. If the present non-canonical conditions required the participants' continuous attention to word order to try to correctly answer the question, CNV activities in the non-canonical conditions were affected by task-driven attention.

The second possible interpretation of the posterior negativities found when encountering non-canonical orders is that they may in fact represent N400 activity, especially considering the temporal and distributional properties of the effects. In previous studies, an enhanced N400 was observed for non-critical words in the final position of an ungrammatical sentence [55, 56]. This N400 was thought to reflect semantic integration difficulty, or a "wrap-up" effect, which is the idea that words following a grammatical violation may become more difficult to integrate semantically. Because the present experiment included only grammatical sentences with the same set of verbs across the four conditions, grammatical and lexical factors were unlikely to induce a semantic integration difficulty for the final verbs in the $CAS_{(+)}/CAN_{(-)}$ and $CAS_{(-)}/CAN_{(-)}$ conditions. Thus, processing a non-canonical word order per se may result in semantic integration difficulty upon encountering the final verb, yielding an N400.

The negative effect for the final verb was more widely observed in anterior sites when non-canonical sentences had non-case-marked NP1s [$CAS_{(-)}/CAN_{(-)}$]. Although the present study did not clarify the sources of the signals for surface potential effects, the ambiguity evoked by non-case-marked NP1s may further increase semantic integration difficulty, leading to the more anterior-distributed negative potential effect which is probably related to an enhanced prefrontal monitoring function.

5. Conclusion

The present study aimed to examine the precise functional and neurophysiological properties underlying discontinuous dependency. The canonical condition beginning with the non-case-marked first word attenuated SAN and yielded P600 only at the third word position. This finding demonstrates that verbal processing is driven by a genuine prediction-based strategy when structural information is not fully available, even if it leads to increased memory and neural resource consumption. That is, a global processing strategy may be fundamentally anchored in a natural language architecture. Future studies should try to elucidate how a global strategy for language comprehension is maintained in clinical populations with not only neurological disorders, but also mental disorders related to goal-related, planning, or attentional functions. We should also investigate how a global strategy in language processing is correlated with general cognitive function.

Acknowledgments

This study was supported by a Grant-in-Aid for Young Scientists (B), The Ministry of Education, Culture, Sports, Science and Technology, Japan (No. 17720116), and a Grant-in-Aid for Challenging Exploratory Research from the Japan Society for the Promotion of Science (No. 15K12884). The authors declare that the research was conducted in the absence of any commercial or financial relationships, which may be interpreted as conflicts of interest. The authors would like to thank all of participants in the present study. We also appreciate Kuniyasu Imanaka, Masami Ishihara, Shinya Kuriki, Kazuyuki Nakagome, Kenji Itoh, Heizo Nakajima, Ken-ichi Takami, Hiroshi Hasegawa, Takeru Honma, Sadayoshi Ogawa, Shiro Ojima, Natsuko Tatsuta, Ayumi Koso, Izumi Kishida, Ryuichiro Hashimoto, and Fumitaka Homae for their help. The first author and colleagues would like to sincerely express their condolences to the late Prof. Hiroko Hagiwara for her valuable contributions to the linguistic and cognitive neuroscience domains of natural language.

REFERENCES

[1] Grodzinsky Y., and Friederici A. D., "Neuroimaging of syntax and syntactic processing," *Curr. Opin. Neurobiol.* **16**, 240–246 (2006). doi:10.1016/j.conb.2006.03.007

[2] Staub, A., Clifton, C., and Frazier, L., "Heavy NP shift is the parser's last resort: evidence from eye movements," *J. Mem. Lang.* **54**, 389–406 (2006). doi:10.1016/j.jml.2005.12.00

[3] Grimshaw, J., Argument Structure, MIT Press, Cambridge, MA, (1990).

[4] Friederici, A. D., and Frish, S., "Verb argument structure processing: the role of verb-specific and argument-specific information," *J. Mem. Lang.* **43**, 476–507 (2000). doi:10.1006/jmla.2000.2709

[5] Chomsky, N., Lectures on Government and Binding, Foris, Dordrecht, (1981).

[6] Chomsky, N., "Derivation by Phase," In Ken Hale: A Life in Language, ed. by M., Kenstowicz, 1-52, MIT Press, Cambridge, MA, (2001).

[7] Clifton, C. J., and Frazier, L., "Comprehending sentences with long-distance dependencies," In *Linguistic Structure in Language Processing*, eds. G. N. Carlson, and M. K. Tanenhaus, Kluwer Academic Publishers, Dordrecht, 273–317 (1989).

[8] Kluender, R., and Kutas, M., "Bridging the gap: evidence from ERPs on the processing unbounded dependencies," *J. Cogn. Neurosci.* **5**, 196–214 (1993). doi:10.1162/jocn.1993.5.2.196

[9] Fiebach, C. J., Schlesewsky, M., and Friederici, A. D., "Separating syntactic memory costs and syntactic integration costs during parsing: the processing of German WH-questions," *J. Mem. Lang.* **47**, 250–272 (2002). doi:10.1016/S0749-596X(02)00004-9

[10] Rosch, E., Mervis, C. B., Gray, W., Johnson, D., and Boyes-Braem, P., "Basic objects in natural categories," *Cogn. Psychol.* **8**, 382–439 (1976).

[11] Caramazza, A. and Zurif, E., "Dissociations of algorithmic and heuristic processes in sentence comprehension: evidence from

aphasia," *Brain Lang.* **3**, 572–582 (1976).
[12] Hagiwara, H., and Caplan, D., "Syntactic comprehension in Japanese aphasics: effects of category and thematic role order," *Brain Lang.* **38**, 159–170 (1990).
[13] Grodzinsky, Y., "A restrictive theory of agrammatic comprehension," *Brain Lang.* **50**, 27–51 (1995). doi:10.1006/brln.1995.1039
[14] Linebarger, M., "Agrammatism as evidence about grammar," *Brain Lang.* **50**, 52–91 (1995). doi:10.1006/brln.1995.1040
[15] Kutas, M., and Hillyard, S. A., "Brain potentials during reading reflect word expectancy and semantic association," *Nature* **307**, 161–163 (1984). doi:10.1038/307161a0
[16] Altmann, G. T. M., "Ambiguity in sentence processing," *Trends Cogn. Sci.* **2**, 146–152 (1998). doi:10.1016/S1364-6613(98)01153-X
[17] Lau, E., Stroud, C., Plesch, S., and Phillips, C., "The role of structural prediction in rapid syntactic analysis," *Brain Lang.* **98**, 74–88 (2006). doi:10.1016/j.bandl.2006.02.003
[18] Summerfield, C., and Egner, T., "Expectation (and attention) in visual cognition," *Trends Cogn. Sci.* **13**, 403–9 (2009). doi:10.1016/j.tics.2009.06.003
[19] Hagiwara, H., "The breakdown of functional categories and the economy of derivation," *Brain Lang.* **50**, 92–116 (1995).
[20] Gibson, E., "Linguistic complexity: locality of syntactic dependencies," *Cognition* **68**, 1–76 (1998). doi:10.1016/S0010-0277(98)00034-1
[21] Gibson, E., "Dependency locality theory: a distance-based theory of linguistic complexity," In Image, Language, Brain, eds. A. Marantz, Y. Miyashita, and W. O'Neil, MIT Press, Cambridge, MA, 95–126 (2000).
[22] Donchin, E., Ritter, W., and McCallum, W. C., "Cognitive psychophysiology: the endogenous components of the ERP," In Event-related Brain Potentials in Man, eds. E. Callaway, P. Tueting, and S. H. Koslow, Academic Press, New York, 349–411 (1978).
[23] King, J. W., and Kutas, M., "Who did what and when?: using word-and clause-level ERPs to monitoring working memory usage in reading," *J. Cogn. Neurosci.* **7**, 376–395 (1995). doi:10.1162/jocn.1995.7.3.376
[24] Rösler, F., Pechmann, T., Streb, J., Roder, B., and Henninghausen, E., "Parsing of sentences in a language with varying word order: word-by-word variations of processing demands are revealed by event-related brain potentials," *J. Mem. Lang.* **38**, 150–176 (1998).
[25] Kaan, E., Harris, A., Gibson, E., and Holcomb, P., "The P600 as an index of syntactic integration difficulty," *Lang. Cogn. Process.* **14**, 631–662 (2000). doi:10.1080/016909600386084
[26] Matzke, M., Mai, H., Nager, W., Russeler, J., and Münte, T., "The costs of freedom: an ERP study of non-canonical sentences," *Clin. Neurophysiol.* **113**, 844–852 (2002). doi:10.1016/S1388-2457(02)00059-7
[27] Felser, C., Clahsen, H., and Münte, T. F., "Storage and integration in the processing of filler-gap dependencies: an ERP study of topicalization and wh-movement in German," *Brain Lang.* **87**, 345–354 (2003). doi:10.1016/S0093-934X(03)00135-4
[28] Ueno, M., and Kluender, R., "Event-related brain indices of Japanese scrambling," *Brain Lang.* **86**, 243–271 (2003). doi:10.1016/S0093-934X(02)00543-6
[29] Bornkessel, I. D., Fiebach, C. J., and Friederici, A. D., "On the cost of syntactic ambiguity in human language comprehension: an individual differences approach," *Cogn. Brain Res.* **21**, 11–21 (2004). doi:10.1016/j.cogbrainres.2004.05.007
[30] Phillips, C., Kazanina, N., and Abada, S. H., "ERP effects of the processing of syntactic long-distance dependencies," *Brain Res. Cogn. Brain Res.* **22**, 407–428 (2005). doi:10.1016/j.cogbrainres.2004.09.012
[31] Hagiwara, H., Soshi, T., Ishihara, M., and Imanaka, K., "A topographical study on the ERP correlates of scrambled word order in Japanese complex sentences," *J. Cogn. Neurosci.* **19**, 175–193 (2007). doi:10.1162/jocn.2007.19.2.175
[32] Friederici, A. D., "Towards a neural basis of auditory sentence processing," *Trends Cogn. Sci.* **6**, 78–84 (2002). doi:10.1016/S1364-6613(00)01839-8
[33] Friederici, A. D., Hahne, A., and Saddy, D., "Distinct neurophysiological patterns reflecting aspects of syntactic complexity and syntactic repair," *J. Psycholinguist. Res.* **31**, 45–63 (2002).
[34] Mangun, G. R., and Hillyard, S. A., "Modulations of sensory-evoked brain potentials indicate changes in perceptual processing during visual-spatial priming," *J. Exp. Psychol. Hum. Percept. Perform.* **17**, 1057–1074 (1991). doi:10.1037/0096-1523.17.4.1057
[35] Friston, K., "Predictive coding, precision and synchrony," *Cogn. Neurosci.* **3**, 238–239 (2012). doi:10.1080/17588928.2012.691277
[36] Hakulinen, A., and Karlsson, F., "Finnish syntax in text," *Nord. J. Linguist.* **3**, 93–129 (1980).
[37] den Ouden, D. B., Saur, D., Mader, W., Schelter, B., Lukic, S., Wali, E., Timmer, J., and Thompson, C. K., "Network modulation during complex syntactic processing," *Neuroimage* **59**, 815–823 (2012). doi:10.1016/j.neuroimage.2011.07.057
[38] Kimball, J., "Seven principles of surface structure parsing in natural language," *Cognition* **2**, 15–47 (1973). doi:10.1016/0010-0277(72)90028-5
[39] Frazier, L., "Sentence processing: a tutorial review," in Attention and Performance XII: The Psychology of Reading, ed. M. Coltheart, Lawrence Erlbaum Associates, Hillsdale, 559–586 (1987).
[40] Phillips, C., and Gibson, E., "On the strength of the local attachment preference," *J. Psycholinguist. Res.* **26**, 323–346 (1997).
[41] Frish, S., Schlesewsky, M., Saddy, D., and Alpermann, A., "The P600 as an indicator of syntactic ambiguity," *Cognition* **85**, 83–92 (2002). doi:10.1016/S0010-0277(02)00126-9
[42] Sakamoto, T., "Processing filler-gap constructions in Japanese: the case of empty subject sentences," In Sentence Processing in East Asian languages, ed. M. Nakayama, CSLI, Stanford, 189–221 (2002).
[43] Kaiser, E., and Trueswell, J. C., "The role of discourse context in the processing of a flexible word-order language," *Cognition* **94**, 113–147 (2004). doi:10.1016/j.cognition.2004.01.002
[44] Fillmore, C. J., "The case for case," In Universals in Linguistic Theory, eds. E. Bach, and R. T. Harms, Holt, Rinehart, and

Winston, New York, 1–88 (1968).
[45] Fodor, J. D., and Inoue, A., “Syntactic features in reanalysis: positive and negative symptoms,” *J. Psycholinguist. Res.* **29**, 25–36 (2000).
[46] Oldfield, R. C., “The Assessment and analysis of handedness: the Edinburgh inventory,” *Neuropsychologia* **9**, 97–113 (1971).
[47] Negishi, N., Maehara, K., and Momose, T., “Handedness and criminals (in Japanese),” *Correctional Medicine*, **39**, 40–47 (1990).
[48] Daneman, M., and Carpenter, P. A., “Individual differences in working memory and reading,” *J. Verb. Learn. Verb. Behav.* **19**, 450–466 (1980).
[49] Osaka, M., Working Memory: The Notepad Inside the Brain (in Japanese), Shinyousya, Tokyo, (2002).
[50] Amano, N., and Kondo, K., Nihongo-no Goitokusei: Lexical Properties of Japanese (in Japanese), Sanseido, Tokyo, (2003).
[51] Klem, G. H., Lüders, H. O., Jasper, H. H., and Elger, C., “The ten-twenty electrode system of the International Federation,” *Electroencephalogr. Clin. Neurophysiol. Suppl.* **52**, 3–6 (1999).
[52] Schlesewsky, M., Bornkessel, I., and Frisch, S., “The neurophysiological basis of word order variations in German,” *Brain Lang.* **86**, 116–128 (2003). doi:10.1016/S0093-934X(02)00540-0
[53] Haarmann, H. J., Cameron, K. A., and Ruchkin, D. S., “Short-term semantic retention during on-line sentence comprehension: brain potential evidence from filler-gap constructions,” *Brain Res. Cogn. Brain Res.* **15**, 178–190 (2003). doi:10.1016/S0926-6410(02)00168-4
[54] Birbaumer, N., Elbert, T., Canavan, A. G. M., and Rockstroh, B., “Slow potentials of the cerebral cortex and behavior,” *Physiol. Rev.* **70**, 1–41 (1990).
[55] Osterhout, L., and Holcomb, P., “Event-related brain potentials elicited by syntactic anomaly,” *J. Mem. Lang.* **31**, 785–806 (1992). doi:10.1016/0749-596X(92)90039-Z
[56] Hagoort, P., Brown, C., and Groothusen, J., “The syntactic positive shift (SPS) as an ERP measure of syntactic processing,” *Lang. Cogn. Process.* **8**, 439–483 (1993). doi:10.1080/01690969308407585