Language, Cognition and Neuroscience

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/plcp21

The role of discourse in long-distance dependency formation

Masataka Yano & Masatoshi Koizumi

To cite this article: Masataka Yano & Masatoshi Koizumi (2021): The role of discourse in long-distance dependency formation, Language, Cognition and Neuroscience, DOI: 10.1080/23273798.2021.1883694

To link to this article: https://doi.org/10.1080/23273798.2021.1883694

LANGUAGE, COGNITION AND NEUROSCIENCE
https://doi.org/10.1080/23273798.2021.1883694

Routledge
Taylor & Francis Group

REGULAR ARTICLE

OPEN ACCESS Check for updates

The role of discourse in long-distance dependency formation

Masataka Yano [a,b] and Masatoshi Koizumi [c,d]

[a]Department of Language Sciences, Tokyo Metropolitan University, Tokyo, Japan; [b]Faculty of Humanities, Kyushu University, Fukuoka, Japan; [c]Department of Linguistics, Graduate School of Arts and Letters, Tohoku University, Sendai, Japan; [d]National Institute for Japanese Language and Linguistics, Tokyo, Japan

ABSTRACT

Sentences with filler-gap dependency are more difficult to process than those without, as reflected by event-related brain potentials (ERPs) such as sustained left anterior negativity (SLAN). The cognitive processes underlying SLAN may support associating a filler with a temporally distant gap in syntactic representation. Alternatively, processing filler-gap dependencies in the absence of a supportive context involves additional discourse processing. The present study conducted an ERP experiment that manipulated syntactic complexity (subject–object–verb [SOV] and object–subject–verb [OSV]) and discourse (the supportive and non-supportive context) in Japanese. The result showed a SLAN in OSV relative to SOV in the non-supportive but not the supportive context, which suggests that the difficulty involved in processing OSV in Japanese is largely due to a pragmatic factor. The present study contributes to a better understanding of how the language-processing system builds long-distance dependency by interacting with the memory system.

ARTICLE HISTORY
Received 25 April 2020
Accepted 20 January 2021

KEYWORDS
Filler-gap dependency; discourse; sustained left anterior negativity (SLAN); Japanese

1. Introduction

In real-time language comprehension, the language-processing system constructs various types of structural relations, including thematic, coreferential, and filler-gap relations. Given the hierarchical nature of natural language, structurally related elements are not necessarily adjacent in a string of successive inputs, thus requiring the language-processing system to form a dependency between the current input and another input, while keeping other relations suspended. The memory demand for this process can be substantial because the temporal distance between structurally related elements is unbounded in principle. Thus, previous studies on human language processing have focused on how the language-processing system associates linearly distant elements by interacting with the memory system.

Among the various types of dependencies, filler-gap dependencies have attracted much academic attention in the literature (Aoshima et al., 2003, 2004; Crain & Fodor, 1985; Frazier & Clifton, 1989; Omaki et al., 2015; Phillips, 2006; Stowe, 1986; Wagers & Phillips, 2014). A filler refers to a dislocated element, such as "*the boy*" in (1), while a gap refers to a filler's original position, which is indicated with the underlining in (1). The subscript index *i* indicates that the two elements refer to the same referent. Since the filler itself does not provide thematic information, the language-processing system needs to carry the filler forward in memory until it finds the original position in order to attain an appropriate thematic interpretation (i.e. the patient of *accused* in (1)).

(1) The policeman saw the boy$_i$ that the crowd at the party ^1 accused __$_i$ ^2 of the crime.

Previous proposals as to how filler-gap dependencies are constructed are broadly categorised into two views, namely the retrieval view and the active maintenance view. The hybrid view combines these two views, as seen below.

1.1. The retrieval view

One of the key findings from various cross-modal lexical priming and probe recognition experiments is that a filler is reactivated at a gap position and associated with the gap (Clahsen & Featherston, 1999; Miyamoto & Takahashi, 2002; Nakano et al., 2002; Nicol & Swinney, 1989). In the cross-modal lexical priming experiment, for example, while participants are listening to a sentence with an object relative clause, as in (1) above, a probe word is visually presented to coincide with the

CONTACT Masataka Yano yanomasa@tmu.ac.jp
Supplemental data for this article can be accessed https://doi.org/10.1080/23273798.2021.1883694

pre-gap position (indicated by ^1) or the gap position (indicated by ^2) (cf. Nicol & Swinney, 1989; Swinney et al., 2000). The probe word is either semantically related or unrelated to the filler. The participants are asked to quickly identify whether the probe word is an actual word in English. Faster reaction times to the semantically-related probe words as compared to the non-related words (i.e. the priming effect) are assumed to reflect the filler's higher activation level. Previous studies have shown a significant priming effect at the gap position (indicating that the filler is active here) but not at the pre-gap position. This result is interpreted to mean that after a filler is processed, its activation decays over time; however, when detecting a gap, the language processing system reactivates the filler by retrieving it from working memory to form a filler-gap dependency. We refer to this interpretation as the retrieval view.

Moreover, some event-related potential (ERP) studies observed a phasic left anterior negativity (LAN) followed by a P600 effect at the gap position. The phasic LAN has been interpreted as an index of the reactivation of a filler (King & Kutas, 1995; Kluender & Kutas, 1993a, 1993b). The P600 has been thought to reflect the syntactic integration of a filler and a gap (Kaan et al., 2000).

The retrieval view fits into a general parsing model called the cue-based parsing model which explains a wide variety of phenomena regarding the interaction between language processing and memory, such as the similarity-based interference effect, the agreement attraction effect, and the forgetting effect (Foraker & McElree, 2011; Lewis et al., 2006; Lewis & Vasishth, 2005; Öztekin et al., 2010; Wagers & McElree, 2013). According to the model, the language-processing system encodes certain information (e.g. the number and cases of noun phrases (NPs)) in content-addressable memory. Afterwards, it accesses the information by retrieving it from memory based on cues, such as "singular" and "nominative case", and then forms a dependency between the current input and the retrieved input. For example, in a sentence like "*It was a boat that the guy who lived by the sea fixed in two sunny days*", some features associated with "*a boat*" are encoded in memory. It is retrieved later on "*fixed*", which signals that an object position can take "*a boat*" as an object (Van Dyke & McElree, 2006). Between encoding and retrieval, information undergoes decay, gradually decreasing the activation level. This model does not assume a specialised system that acts as a buffer (contra Baddeley & Hitch, 1974; Wanner & Maratsos, 1978).[1] Thus, the filler-gap dependency is treated in the same way as other dependencies, such as thematic subject–verb dependency.

1.2. The active maintenance view

There exists conflicting evidence supporting another view wherein a filler is kept active in memory until a gap is detected. In ERP experiments, a sustained left anterior negativity (SLAN) has been observed between a filler and a gap in object wh-questions (in comparison to subject wh-questions and complement clauses without a gap) (Fiebach et al., 2001; Phillips et al., 2005), post-nominal object relative clauses (in comparison to subject relative clauses) (King & Kutas, 1995; Müller et al., 1997), and scrambled sentences (in comparison to syntactically basic sentences) (Hagiwara et al., 2007; Matzke et al., 2002; Ueno & Kluender, 2003). The SLAN has been proposed to reflect the (syntactic) working memory load that is needed to actively maintain a filler (Hagiwara et al., 2007; King & Kutas, 1995; Matzke et al., 2002; Müller et al., 1997; Phillips et al., 2005).[2] The SLAN has been thought to be similar to the slow negative potential reported by classic monkey studies (Batuev et al., 1985; Funahashi et al., 1989). Monkeys have showed a sustained activity in the prefrontal cortex during delayed response tasks in which they were required to remember task-relevant stimulus information for a short interval and execute a trained reaction on cue.

Functional magnetic resonance imaging (fMRI) studies also found increased activity during filler-gap dependency formation (Fiebach et al., 2005; Kim et al., 2009; Kinno et al., 2008; Rogalsky et al., 2008; Rogalsky et al., 2015; Santi & Grodzinsky, 2007, 2010). The increased activity is located in the left inferior frontal region (LIFG), such as the pars triangularis and the pars opercularis. According to Santi and Grodzinsky (2007), the LIFG activation intensifies as the filler-gap distance increases (e.g. short: The mailman and the mother of Jim love the woman$_i$ who Kate pinched ___$_i$ vs. long: Kate loves the women$_i$ who the mailman and the mother of Jim pinched ___$_i$). This trend was not observed in the long-distance binding relationship, in which the language-processing system needs to retrieve an appropriate antecedent for a reflexive (e.g. short: The sister of Kim assumes that Anne loves the mailman$_i$ who pinched himself$_i$. vs. long: Anne assumes that the mailman$_i$ who loves the sister of Kim pinched himself$_i$). This suggests that LIFG activity contributes to syntactic memory rather than to memory retrieval in order to keep information active (cf. Grodzinsky & Santi, 2008; Santi &

Grodzinsky, 2007; but see also Matchin et al., 2014). These findings support the view that the displaced filler is kept active while the filler-gap dependency is being processed (i.e. the active maintenance view).

Furthermore, Frazier and Flores d'Arcais (1989) interpreted behavioural study results using self-paced reading methods in alignment with the active maintenance view (Frazier, 1987; Frazier & Clifton, 1989; Frazier & Flores d'Arcais, 1989). A reading slow-down was observed when positing a gap at a potentially available position nearest a filler (e.g. a subject position, a parasitic gap position inside the subject, and an object position) that was blocked due to an overt NP or semantic implausibility (Lee, 2004; Phillips, 2006; Stowe, 1986). For instance, Stowe (1986) has reported that readers showed a slow down at "*us*" in (2) because, before confirming that the position was occupied by another element they expected "*who*" to originate from the object position (i.e. the active filler strategy):

(2) My brother wanted to know who$_i$ Ruth will bring us home to __$_i$ at Christmas.

Frazier and Flores d'Arcais (1989) argued that the filler is unlikely to be inactive during the search for a gap position.

1.3. The hybrid view and research questions

These two views seem contradictory. Why does the language-processing system reactivate a filler from memory even though it actively maintains the filler? To reconcile this issue, Wagers and Phillips (2014) have proposed a hybrid view on the basis of a series of self-paced reading experiments. According to the study, the language-processing system actively holds coarse information, such as syntactic category (e.g. NP), that helps find a gap, but it releases more fine-grained information, such as lexico-semantic information, that is later reactivated when detecting a gap. This account successfully combines the active maintenance view that is founded on syntax-based paradigms, like the filled-gap effect in (2), with the retrieval view, which is founded on semantic-based paradigms, like the priming effect in (1).

Assuming that the account is on the right track, what do ERPs reveal about the time-course of filler-gap dependency formation? A LAN effect at the gap position is considered to be an ERP correlate of retrieving semantic information, and a following P600 reflects the association of the reactivated filler with a gap. What, then, does SLAN reflect? A straightforward interpretation under this view would be that it reflects the memory process of keeping syntactic information active. However, Yano and Koizumi (2018) recently challenged this idea based on their observation of an interaction between filler-gap dependency and discourse, as explained in the next section.

1.4. Discourse effects on the processing of filler-gap dependency

Syntactic complexity correlates with discourse factors such as topichood, focus, and givenness (Aissen, 1992; Birner & Ward, 2009; Kuno, 1987). A language's basic word order is a default option for describing an event and occurs in a wide range of contexts. In contrast, a more syntactically complex word order that is derived through movement is a marked choice, and its use must be well-justified by the discourse factors. The literature on language processing references previous studies that have demonstrated that contextual support facilitates reading times for non-basic word orders in accordance with this theoretical claim (Clifton & Frazier, 2004; Grodner et al., 2005; Imamura et al., 2016; Kaiser & Trueswell, 2004; Koizumi & Imamura, 2017; Meng et al., 1999; Sekerina, 2003). For example, Kaiser and Trueswell (2004) conducted a self-paced reading experiment in Finnish, using two types of context (supportive/non-supportive) and word orders (syntactically basic subject-verb-object [SVO] / non-basic object-verb-subject [OVS]), as shown in (3) and (4), respectively. The supportive context in (3a) referred to the object of the target sentences in (4b) to license a felicitous use of OVS, in which the object must be discourse-old information in Finnish; the non-supportive context in (3b) did not. The two factors showed a significant interaction with regard to the second argument (which is underlined),

(3) Preceding context

Lotta	etsi	eilen	sieniä	metsässä.	Hän	huomasi
Lotta	looked-for	yesterday	mushrooms	forest-in	She-NOM	noticed
heinikossa	(a)jäniksen /(b)hiiren	joka	liikkui	varovasti	eteenpäin.	
grass-in	hare-ACC/mouse-ACC	that	was.moving	carefully	forward.	

'Lotta looked for mushrooms in the forest yesterday. She noticed {(a) a hare /(b) a mouse} moving forward carefully in the grass.'

(4) a. SVO

Hiiri	seurasi	jänistä	ja	linnut	lauloivat.
mouse-NOM	followed	hare-PART	and	birds	were.singing.

b. OVS

Jänistä	seurasi	hiiri	ja	linnut	lauloivat.
hare-PART	followed	mouse-NOM	and	birds	were.singing.

'The mouse followed the hare and birds were singing.'

with a longer reading time in OVS than in SVO only in the non-supportive context.

The discourse-givenness effect is also at play in Japanese. In Japanese, object–subject–verb (OSV) is derived from subject–object–verb (SOV) by fronting an O over an S (Saito, 1985, see Appendix A for more detail). To order the discourse information coherently, OSV is used when O refers to information that is discourse-older than the information to which S refers (Kuno, 1987, Information Flow Principle). In all other cases, a syntactically simpler SOV is preferred over an OSV. According to corpus studies, the O in OSV was discourse-old information in 81% of OSV occurrences in Japanese (Imamura, 2014; Imamura, 2015; Imamura & Koizumi, 2011). Furthermore, Ferreira and Yoshita (2003) showed that when native Japanese speakers were asked to recall a ditransitive sentence with a new-given order, they produced the sentence in the given-new order by fronting a discourse-given argument over a discourse-new argument. These observations support the assertion that givenness plays a crucial role in motivating scrambling in Japanese.

Building on these observations, Yano and Koizumi (2018) conducted an ERP experiment similar to Kaiser and Trueswell's (2004) experiment. They manipulated the context and word orders in Japanese, as given in (5) and (6). They found a larger SLAN from a filler (i.e. O) to a gap in OSV in comparison to SOV, but only in the non-supportive context. A P600 effect was observed at the S of the OSV in the non-supportive context but not in the supportive context. These results suggest that the difficulty involved in processing OSV is largely due to discourse factors, not syntactic representational complexities.

1.5. Discourse effects on syntactic complexity

The interpretation of Yano and Koizumi's (2018) experiment requires a caveat (Yano, 2019). It is known that a topic can attain a patient/theme interpretation without forming a filler-gap dependency, as evidenced by the acceptability of (7) in which a putative filler, "*sono-e-wa*", crosses a relative clause island (Kuno, 1987). From the language-processing perspective, it is possible that the language-processing system interprets a discourse-given O as a topic and does not create a filler-gap dependency when OSV is used in a felicitous context. If this is the case, the lack of both SLAN and P600 should count as the lack of filler-gap dependency in OSV, rather than the possibility that the discourse ameliorates filler-gap dependency processing. Thus, the present study uses a negative polarity item (NPI), "*shika*", that forces the language-processing system to reconstruct a filler in its original position.

1.6. NP-shika in Japanese

"*Shika*" (nothing/nobody but) is an NPI that is attached to NPs in Japanese.[3] NP-shika appears with a negation, as in (8a), but cannot appear with an affirmative predicate, as in (8b). It has generally been assumed that to be licensed, NP-shika must be generated within a c-commanding domain of negation (Kato, 1985, 1994). Once the licensing relation is established, it becomes possible to reorder NP-shika, as in (8c).

(5) Context:

Kooban-ni	(a) Yoshida-san-ga /(b) Kimura-san-ga	imasu.
police.box-in	Yoshida-Mx.-NOM / Kimura-Mx.-NOM	be

'(a) Mx. Yoshida /(b) Mx. Kimura is in the police box.'

(6) a. SOV

Yoshida-san-ga	kinoo-no	yoru	Kimura-san-o	yurushita	rashii.
Yoshida-Mx.-NOM	yesterday-GEN	night	Kimura-Mx.-ACC	forgave	seem

'It seems that Mx. Yoshida forgave Mx. Kimura last night.'

b. OSV

Kimura-san-o	kinoo-no	yoru	Yoshida-san-ga	yurushita	rashii.
Kimura-Mx.-ACC	yesterday-GEN	night	Yoshida-Mx.-NOM	forgave	seem

'It seems that Mx. Yoshida forgave Mx. Kimura last night.'

(7)

Sono-e$_i$-wa	Taro-ga	$[_{NP}[_{TP}$	e_i	kaita]	hito]-o	yoku	shitteiru.
that.picture-TOP	Taro-NOM		e	drew	person-ACC	well	know

'Speaking of that painting, Taro knows the person who drew (it).'

(8)	a.	Taro-ga	sushi-shika	tabe-nakat-ta.[4]
		Taro-NOM	sushi-shika	eat-NEG-PST
		'Taro ate nothing but sushi.'		
	b. *	Taro-ga	sushi-shika	tabe-ta.
		Taro-NOM	sushi-shika	eat-PST
	c.	Sushi-shika	Taro-ga	tabe-nakat-ta.
		sushi-shika	Taro-NOM	eat-NEG-PST
		'Taro ate nothing but sushi.'		

In order to arrive at a correct interpretation, the NP-shika must be interpreted in the vP domain, even if it moves out of vP. Such evidence comes from scope interpretation (Kataoka, 2010). First, consider the sentence in (9). A quantifier phrase (QP), "*more than five books*", is scopally ambiguous in relation to the negation in the sentence without "*-shika*" in the subject position. It can take either a narrower or wider scope over the negation. This contrasts with the sentences in (10), in which the subject is marked with "*-shika*". The QP "*more than five books*" cannot take a wider scope over the negation because the NP-shika is necessarily interpreted within vP, and thus the structurally lower QP cannot escape negation.

(9)	Sono-gakusei-ga	go-satsu-izyo-no-hon-o	yom-anai.
	the-student-NOM	five-classfier-more.than-GEN-book-ACC	read-NEG

(i) NEG > QP: There are not five or more books which the student reads.
(ii) QP > NEG: There are five or more books which the student does not read.

(10)	Sono-gakusei-shika	go-satsu-izyo-no-hon-o	yom-anai.
	the-student-shika	five-classfier-more.than-GEN-book-ACC	read-NEG

(i) NEG > QP: All but that student do not read five or more books.
(ii) *QP > NEG: There are five or more books that all but that student do not read.

To summarise, the NP-shika does not undergo raising to achieve proper licensing. When it happens to move beyond the NegP through another operation, namely scrambling, the operation must be undone to facilitate the interpretation. This property of NP-shika offers good ground for testing the current hypotheses. From the perspective of real-time language processing, when object-shika appears in the sentence-initial position, the language-processing system is supposed to reconstruct a filler in its original position in order to check whether it is correctly licensed by a negation, regardless of discourse. If SLAN reflects an active maintenance cost, this process should elicit the SLAN effect. Otherwise, it would not necessarily appear when the object-shika is scrambled to the sentence-initial position. To test these possibilities, the present study conducted an ERP experiment that attempts to clarify how the language-processing system associates a filler with a gap by interacting with the memory system. We demonstrate that the discourse affects how the language-processing system processes the dependency between a filler and gap.

2. The ERP experiment

2.1. Participants

Eighteen native Japanese speakers participated in the present experiment (nine females and nine males, mean age = 21.4, standard deviation (SD) of age = 2.4). They are undergraduate or graduate students at Kyushu University. All participants were classified as right-handed, based on the Edinburgh handedness inventory (mean = 95.3, range: 70–100, SD = 7.6) (Oldfield, 1971).[5] The participants had normal or corrected-to-normal vision and reported no history of language/neurological disorders. This experiment has been approved by the Ethics Committee of the Department of Linguistics, Kyushu University. Written informed consent was obtained from all participants prior to the experiment, and they were compensated for their participation.

2.2. Stimuli

The sentences (11) and (12) are an example set of contexts and target sentences. WORD ORDER (SOV/OSV) and GIVENNESS (new-given order/given-new order) were manipulated to create four conditions.

The givenness of arguments in (12) was manipulated by presenting contexts, such as in (11), in which one of the two persons referred to either the target sentences' S or O. Two persons were introduced, since at least two persons must exist in the discourse to render the use of NP-shika appropriate. The order of a person to be introduced in a target sentence (i.e. Mx. Kitamura and Mx. Yamada or the reverse order) was counterbalanced across trials to avoid participants' being able to predict who would be the first or second person to appear in a target sentence. As discussed above, SOV can be used in a wider range of contexts, allowing for given-new and new-given orders. In contrast, OSV requires the O to be discourse-given. Thus, the context in (11a) renders OSV felicitous (i.e. the supportive context), whereas the context in (11b) does not provide any support for it (i.e. the non-supportive context).

The NPs of the contexts and target sentences were common family names, with no bias with regard to particular thematic roles ("*-san*" is an honorific term that can refer to males and females). Henceforth, we refer to the target sentences' first noun phrase (i.e. "Aoki-san-ga" Mx.

(11)	Context:						
	(a)	Kitamura-san-to	Yamada-san-ga	kaigishitsu-ni	imas-ita.		
		Kitamura-Mx.-and	Yamada-Mx.-NOM	meeting.room-LOC	be-PST		
		'Mx. Kitamura and Mx. Yamada were in the meeting room.'					
	(b)	Aoki-san-to	Yamada-san-ga	kaigishitsu-ni	imas-ita.		
		Aoki-Mx.-and	Yamada-Mx.-NOM	meeting.room-LOC	be-PST		
		'Mx. Aoki and Mx. Yamada were in the meeting room.'					
(12)	Target sentences:						
	(a) SOV:						
		Aoki-san-ga	senshu-no	getsuyobi	Kitamura-san-ni-shika	aisatsushi-nakat-ta	rashii
		Aoki-Mx.-NOM	last.week-GEN	Monday	Kitamura-Mx.-DAT-NPI	greet-NEG-PST	seem
		'Mx. Aoki seemed to greet only Mx. Kitamura last Monday.'					
	(b) OSV:						
		Kitamura-san-ni-shika	senshu-no	getsuyobi	Aoki-san-ga	aisatsushi-nakat-ta	rashii
		Kitamura-Mx.-DAT-NPI	last.week-GEN	Monday	Aoki-Mx.-NOM	greet-NEG-PST	Seem
		'Mx. Aoki seemed to greet only Mx. Kitamura last Monday.'					

Aoki-NOM in the SOV condition and "Kitamura-san-ni-shika" Mx. Kitamura-DAT-NPI in the OSV condition) as NP1, and the second NP (i.e. "Kitamura-san-ni-shika" Mx. Kitamura-DAT-NPI in the SOV condition and "Aoki-san-ga" Mx. Aoki-NOM in the OSV condition) as NP2. Temporal adverbs occurred between the S and the O to observe a SLAN. The type of verbs used in the present experiment differed from that in Yano and Koizumi's (2018) study. In Japanese, mono-transitive verbs take either an accusative or a dative case for objects. In the accusative-taking verbs that Yano and Koizumi (2018) used, an accusative case must drop when an NP attaches *"shika"* (*NP-ACC-shika, ✓NP-shika). This should be problematic because non-case-marked NPs induce a temporal subject–object ambiguity at the sentence-initial position. Therefore, we used dative-taking verbs, which can take on the form *"NP-DAT-shika"*.

One hundred and twenty sets of experimental stimuli were distributed on four lists, following a Latin square design, so that each participant read only one sentence from the same set. The lists were counterbalanced across the participants.

2.3 Procedure

Sentences were presented in the centre of the monitor in random order, using Presentation ver. 21.1 (Neurobehavioral Systems). At the beginning of a trial, a fixation was presented for 700 ms, followed by a blank screen for 300 ms. A lead-in context was presented in its entirety for 3000 ms, followed by an inter-stimulus interval (ISI) of 200 ms. Then, each phrase of the target sentences was presented for 700 ms with 200 ms ISI. The participants were instructed not to blink while reading the target sentences. At the end of a trial, a comprehension question was presented to check whether the participants had paid enough attention to the sentences' contents. The participants were required to answer questions by pressing either the YES or NO button on the response pad (Cedrus, RB-740). The number of yes and no responses was balanced across trials.

2.4. Electrophysiological recordings

EEGs were recorded from nineteen Ag electrodes (QuickAmp, Brain Products) located at Fp1/2, F3/4, C3/4, P3/4, O1/2, F7/8, T7/8, P7/8, Fz, Cz, and Pz, according to the international 10–20 system (Jasper, 1958). Additional electrodes were placed to the left of and below the left eye to monitor horizontal and vertical eye movements. The reference was set to the average of all electrodes online, and EEGs were re-referenced to the average value of the earlobes offline. The impedances of all electrodes were maintained at less than 10 kΩ during the experiment. The EEGs were amplified with a bandpass of DC to 200 Hz, digitised at 1000 Hz.

2.5. Electrophysiological data analysis

Independent component analysis (ICA) was performed to ensure that SLAN is not affected by EOG artefacts in the following procedures: (i) EEGs were down-sampled to 500 Hz to facilitate ICA. (ii) The band-pass filter was applied at 1–40 Hz. (iii) ICA was performed for continuous EEGs, using the adaptive mixture ICA (AMICA) toolbox equipped in EEGLAB (Palmer & Makeig, 2008). (iv) EEG epoching was performed from −500 to 2700 ms, relative to the onset of the target sentences' NP1 (i.e. "Aoki-san-ga" Mx. Aoki-NOM in the SOV condition and "Kitamura-san-ni-shika" Mx. Kitamura-DAT-NPI in the OSV condition). (v) EOG artefacts were identified using the ICLabel toolbox (Pion-Tonachini et al., 2019). This toolbox assigns the probability of eye, muscle, heart, line noise, channel noise, and others to each independent component. A component was assumed to be an EOG artefact and removed when the probability of "eye" exceeded 0.90. (vi) The baseline was set to −100–0 ms, and the EEGs were re-referenced to the average value of the earlobe electrodes. (vii) They were averaged across trials in each channel and condition. Trials with large artefacts (exceeding ±150 μV) were automatically removed.[6] EEGs were filtered using a 10 Hz low-pass filter only for plotting. All processing was conducted using EEGLAB (Delorme & Makeig, 2004) and ERPLAB (Lopez-Calderon & Luck, 2014).

The present study focuses on two ERP effects, namely SLAN and P600. Following previous studies on the processing of long-distance dependency, two types of analyses were conducted to assess these (Fiebach et al., 2001; King & Kutas, 1995; Phillips et al., 2005; Ueno & Kluender, 2003, 2009). The multi-word analysis examined SLAN from the onset of NP1 to the onset of NP2.[7] The SLAN was expected to appear after lexical access to NP1 was completed (i.e. approximately 300–500 ms). Hence, for NP1, SLAN's presence was examined to compare the mean amplitude of 500–900 ms. The time-window of 300–500 ms was also tested to examine the priming effect exerted by the context sentence. For the following two adverbs, the SLAN was quantified by calculating the mean amplitude from 100 ms after the onset of each region to the end of the epoch (100–900 ms) (Lau & Liao, 2018; Phillips et al., 2005; Yano & Koizumi, 2018). For each word, the first 100 ms time-window was excluded from the analyses because basic perceptual processing is performed during this time-window, and thus ERPs seem to offer little information about dependency formation (cf. Lau & Liao, 2018, p. 638). The baseline correction was not applied at the onset of adverbs. The single-word analysis examined a P600 at NP2. The ERPs were quantified by calculating the mean amplitude for each participant using two time-windows: 300–500 ms and 500–700 ms.[8]

All statistical analyses were conducted separately at the midline (Fz, Cz, and Pz), lateral (F3, F4, C3, C4, P3, and P4), and temporal (Fp1, Fp2, F7, F8, T7, T8, P7, T8, O1, and O2) arrays. The midline analysis consisted of repeated measures ANOVA, with three within-group factors: Word Order (SOV/OSV) × Givenness (given-new/new-given) × Anteriority. The lateral and temporal analyses involved four within-group factors: Word Order × Givenness × Hemisphere (left/right) × Anteriority. When an interaction occurred between Word Order × Givenness, the effect of Word Order was examined at each level of Givenness and that of Givenness was examined at each level of Word Order. When Word Order and/or Givenness interacted with topographic factors (Anteriority and/or Hemisphere), subsequent analyses were conducted at each topographic factor level (e.g. front, central, and posterior). The Greenhouse–Geisser correction was applied for all effects involving more than one degree of freedom (Greenhouse & Geisser, 1959). Given that the main effect of Anteriority/Hemisphere and their interaction, which does not involve experimental conditions, were of no interest, we did not report them below for simplicity of exposition.

2.6. Prediction

Previous studies have proposed that SLAN is an index of the working memory load for actively holding a filler in working memory (Hagiwara et al., 2007; King & Kutas, 1995; Kluender & Kutas, 1993a; Matzke et al., 2002; Müller et al., 1997; Phillips et al., 2005). If this is the case, we expect that SLAN will not be modulated by Givenness. On the other hand, if it reflects a discourse-level processing cost to accommodate the discourse requirement encoded by non-basic word orders, we predict that a felicitous context will ameliorate it, leading to a lack of SLAN.

P600 has been observed at the gap position in filler-gap dependency and proposed to reflect the syntactic processing difficulty of associating a filler with its original position (Kaan et al., 2000). We expect no givenness effect if the P600 reflects a syntactic integration difficulty. On the other hand, if it relates to a discourse-level processing cost, we predict a P600 only in the infelicitous context.

In addition to these ERP effects, a larger N400 is expected for new NPs at NP1 (the S of SOV and the O of OSV) and NP2 (the O of SOV and the S of OSV), since N400 has been known to attenuate due to priming (Kutas & Federmeier, 2011). Although this effect is reported together with SLAN and P600, it is of no interest for the present purpose.

An anonymous reviewer raised important questions as to why the SLAN extends over several words instead of being localised at the O and what the P600 reflects if the SLAN is associated with presupposition accommodation. In accommodating an unsatisfied presupposition, at least two distinct processes are supposed to be necessary. First, when a felicitous context is not provided, comprehenders need to either execute an extensive memory search to find an appropriate context that renders the use of the OSV sentence felicitous or come up with an appropriate context on their own. This process might not be completed within a short time interval after encountering an O, and in such cases, it extends over several words downstream, which is reflected by a long-lasting SLAN. Second, comprehenders subsequently need to integrate the implicit context with an OSV sentence to attain a coherent discourse representation. This process occurs when a non-canonical OSV word order is determined (i.e. at the point of a nominative S), leading to a P600 effect.

Burkhardt (2006) and Burkhardt (2007) have been investigated the cases in which comprehenders need to accommodate unsatisfied presupposition. Burkhardt (2006) reported a P600 effect in German when definite articles' existential presuppositions (cf. Heim, 1982) were not fully satisfied (*Tobias visited a concert in Berlin. He said that the conductor was very impressive.*) and when they were not satisfied at all (*Tobias talked to Nina. He said that the conductor …*), compared to when they were satisfied by the context (*Tobias visited a conductor in Berlin. He said that the conductor …*). This observation suggests that additional discourse processing elicits a P600 effect.

3. Results

3.1. Behavioural data

The mean accuracy of the comprehension questions was 83.2% ($S_{NEW}O_{GIVEN}V$: 85.3%, $S_{GIVEN}O_{NEW}V$: 85.4%, $O_{NEW}S_{GIVEN}V$: 80.0%, $O_{GIVEN}S_{NEW}V$: 80.7%). The logistic mixed-effect model showed a significant effect of Word Order, indicating that OSV was more difficult to understand than SOV ($\beta = -0.42$, $z = -3.50$, $p < 0.01$). The effect of Givenness and the interaction were not significant ($p > 0.10$).[9]

3.2. Electrophysiological data

3.2.1. Multi-word analysis

Figure 1 shows the grand average ERPs from the onset of NP1 to that of NP2 of the target sentence. A visual inspection suggests that $O_{NEW}S_{GIVEN}V$ (i.e. infelicitous use of OSV) diverged from the other three conditions at Fp1. As shown in the isovoltage map (bottom right), the sustained negativity was localised in the left anterior region.

3.2.1.1. NP1. On NP1, the main effect of Givenness was significant at all arrays in the time-window of 300–500 ms, showing that the discourse-given NPs attenuated an N400 amplitude compared to the discourse-new NPs (Figure 2, Table 1).

In the 500–900 ms time-window, the interaction of Word Order × Givenness × Anteriority was marginally significant at the lateral array. Planned comparison at OSV revealed a significant effect of Givenness at F3/4 ($F(1, 17) = 5.12$, $p < 0.05$). No significant effect of Givenness was found at SOV. The interaction of Givenness × Hemisphere was marginally significant at the temporal array, but a significant effect of interest was not observed in planned comparison (see Appendix B for more detail).

3.2.1.2. The first adverb. In Adv1 (corresponding to 1000–1800 ms in Figure 1), the interaction of Word Order × Givenness × Anteriority was marginally significant at the lateral array. The four-way interaction was also marginally significant at the temporal array. Planned comparison at each level of Givenness found a significant effect at Fp1/2 only at OSV ($F(1, 17) = 6.35$, $p < 0.05$). This significant effect means that the ERP of $O_{NEW}S_{GIVEN}V$ was more negative than that of $O_{GIVEN}S_{NEW}V$.

3.2.1.3. The second adverb. In Adv2 (corresponding to 1900–2700 ms in Figure 1), the four-way interaction reached a significant level. Subsequent analyses at each level of Word Order revealed a significant effect at Fp1 at OSV ($F(1, 17) = 6.55$, $p < 0.05$). No significant difference was found at SOV.

In summary, $O_{NEW}S_{GIVEN}V$ showed a SLAN from the 500 ms post-NP1 onset to the Adv2, mainly at Fp1. No comparable negativity was observed when OSV was felicitously used (i.e. $O_{GIVEN}S_{NEW}V$).

3.2.2. Single-word analysis

Figure 3 shows the grand average ERPs at the target sentence's NP2.

In the 300–500 ms time-window, the main effect of Givenness was significant at the midline and lateral arrays, showing a priming effect (Table 2). In the 500–700 ms time-window, the main effect of Givenness was marginally significant at the midline array, reflecting that the N400 effect continued to this time-window. The significant three-way interaction of WORD ORDER, ANTERIORITY, and HEMISPHERE at the lateral array reflects a small positivity observed at the F4 ($F(1, 17) = 1.64$, $p = 0.08$) and C4 ($F(1, 17) = 1.32$, $p = 0.07$).

To summarise the results at NP2, the discourse-given NPs showed a reduced N400 amplitude in comparison with the discourse-new NPs. In both types of OSV sentences, there was no P600 effect observed.

4. Discussion

The present experiment examined whether the difficulty processing filler-gap dependency reflects an infelicitous use of OSV, rather than the cost of a syntactically complex representation. The result showed an interaction between word order and givenness. $O_{NEW}S_{GIVEN}V$ elicited a SLAN from the O to the S. However, $O_{GIVEN}S_{NEW}V$ did not exhibit a SLAN. At NP2, OSV did not elicit a P600 effect compared to SOV. These results are discussed in the following sections.

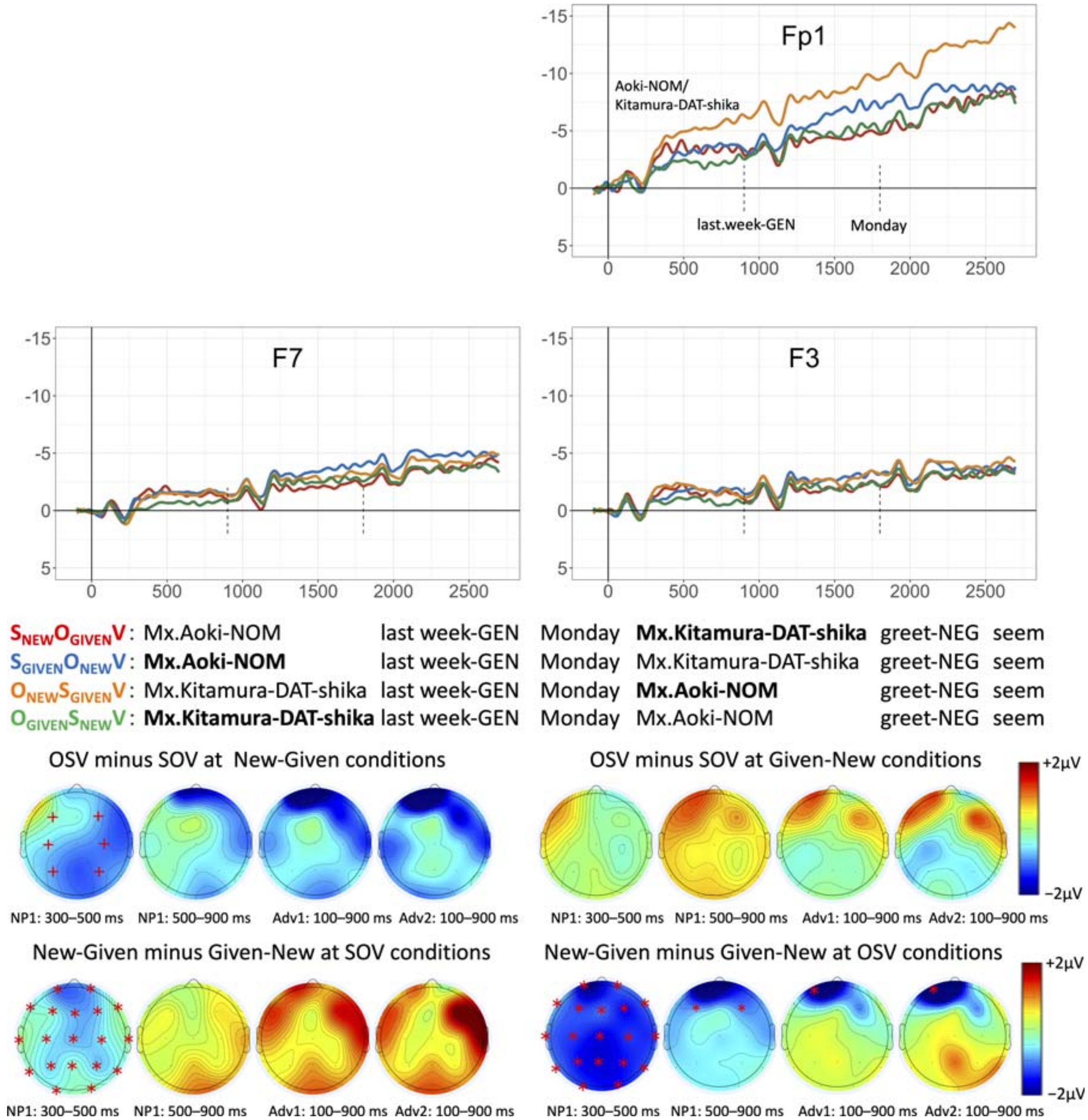

Figure 1. Grand average ERPs from the onset of NP1 (Mx. Aoki-NOM of SOV and Mx. Kitamura-DAT-shika of OSV) to the onset of NP2 (Mx. Kitamura-DAT-shika of SOV and Mx. Aoki-NOM of OSV) (Boldface in the legend indicates discourse-given NPs). The asterisks and plus marks indicate significant and marginally significant effects, respectively.

4.1. The role of SLAN in filler-gap dependency formation

During the 300–500 ms time-window of NP1 and NP2, the N400 amplitude for the discourse-given NPs attenuated relative to the discourse-new NPs, suggesting that the repetition of the same NPs facilitated lexical processing, according to the standard interpretation of N400 (Kutas & Federmeier, 2000, 2011; Lau et al., 2008; Lau et al., 2009). Neither the main effect of Word Order nor the interaction of Word Order × Givenness was significant, suggesting that only word-level processing was performed during this time-window.

During the subsequent time-windows, $O_{NEW}S_{GIVEN}V$ started to diverge from $O_{GIVEN}S_{NEW}V$. We observed a SLAN effect in OSV, but only when its use was not strongly motivated by the context. In other words, $O_{GIVEN}S_{NEW}V$ did not elicit a SLAN effect, even though it involved a filler-gap dependency. Although it is possible that the discourse-new NP was more costly to process than the discourse-given NP, the SLAN cannot solely be attributed to the givenness of NP1 itself. This is because in the SOV comparison, the discourse-new NP did not show a comparable SLAN relative to the discourse-given NP. For the same reason, the difference in the number of NPs encountered by the time of the NP1's

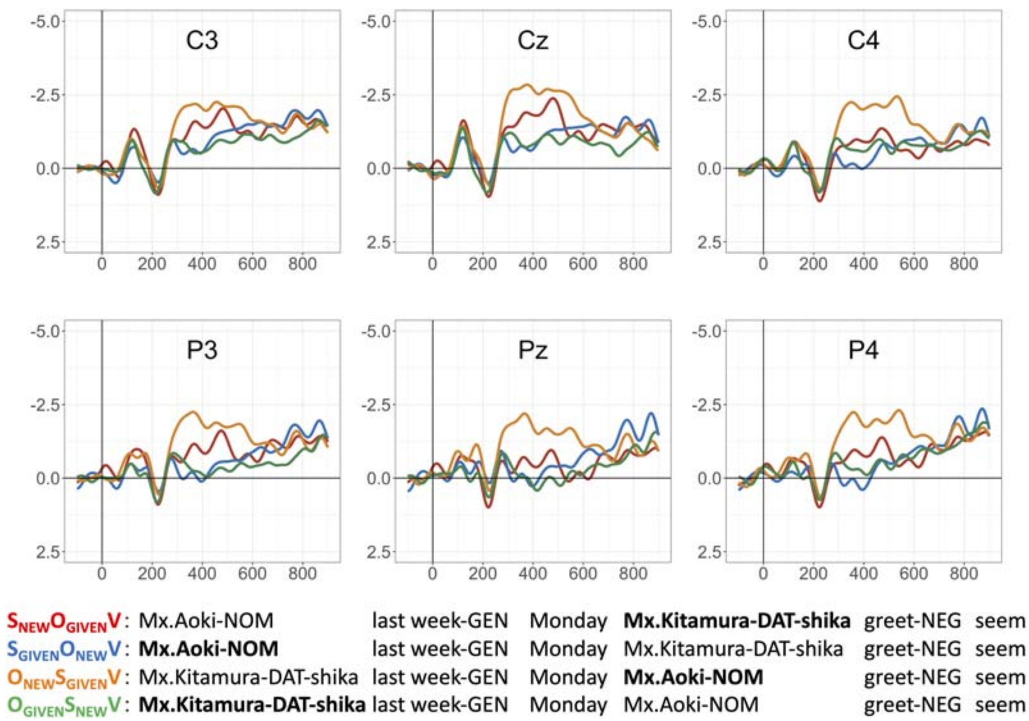

Figure 2. Grand average ERPs at NP1 (Mx. Aoki-NOM of SOV and Mx. Aoki-NOM of OSV) (Boldface in the legend indicates discourse-given NPs).

appearance cannot account for the SLAN. Under the new-given conditions, the participants read three NPs (two NPs in the context and a discourse-new NP in the target sentence) by the NP1, whereas they read two NPs (two NPs in the context, one of which appeared in the target sentence) under the given-new conditions. Since these NPs had to be remembered to perform the task correctly, one might think that this difference affects the working memory load, eliciting a larger SLAN in $O_{NEW}S_{GIVEN}V$. This idea should predict a SLAN effect in $S_{NEW}O_{GIVEN}V$ relative to $S_{GIVEN}O_{NEW}V$, which was not consistent with the present results. The SLAN effect should be explained by the combination of word order and givenness.

The observation of the SLAN effect in $O_{NEW}S_{GIVEN}V$ is consistent with the results of previous studies in Japanese, in which no context was provided, and thus, the presupposition for the use of OSV was not satisfied (Hagiwara et al., 2007; Ueno & Kluender, 2003). The prevalent interpretation of SLAN is that it supports the assertation that the language-processing system holds a filler actively in working memory until a gap position becomes available. However, this interpretation contradicts the result of $O_{GIVEN}S_{NEW}V$, which did not show a SLAN. The absence of a SLAN cannot be explained by the absence of the O's movement status, unlike in Yano and Koizumi's study (2018). Since the language-processing system needs to check whether "*shika*" is legitimately licensed by the negation and attain a correct interpretation of the filler as a patient/theme, it should keep a filler, NP-DAT-shika, in memory until it finds a gap position, regardless of the discourse-givenness of the NP. A possible interpretation of these results is to assume that the discourse ameliorates the process cost of keeping a filler in working memory. However, this interpretation faces the problem of why the pragmatic factor can play a role in syntactic dependency formation. As Yano and Koizumi (2018) discussed, the SLAN has been observed for a semantically vacuous filler, such as "*wer*" (who-ACC) in German (Thomas asks himself, who-ACC on Tuesday afternoon after the accident the doctor ___ called has, Fiebach et al., 2002). Furthermore, Wagers and Phillips (2014) have proposed that a filler's syntactic categorical information (e.g. PP and NP) is actively maintained between the filler and the gap, while lexico-semantic information is not. Lexico-semantic information is quickly displaced from working memory, but it is retrieved later, upon finding a gap. Given these findings, it is possible to suppose that SLAN reflects the active maintenance of a filler's syntactic information, rather than lexico-semantic information. In this interpretation, however, it is unclear how such a process is alleviated by the discourse due to the language-processing system's need to remember a filler's syntactic category, regardless of the context.

Table 1. Statistical results of the multi-word analysis

	NP1: 300–500 ms			NP1: 500–900 ms			Adv1: 100–900 ms			Adv2: 100–900 ms		
	Midline	Lateral	Temporal	Midline	Lateral	Temporal	Midline	Lateral	Temporal	Midline	Lateral	Temporal
Word Order (WO)	2.47	3.03	1.50	1.74	0.76	< 0.01	0.08	0.69	2.40	0.12	0.66	1.96
Givenness (G)	21.52 **	22.77 **	10.26 **	0.96	0.87	0.61	0.05	0.05	0.24	0.12	0.05	0.19
WO × Anteriority (Ant)	0.16	2.37	0.13	0.28	1.82	1.31	0.25	0.76	1.06	0.01	0.99	1.08
G × Ant	0.90	0.63	0.87	1.34	0.92	1.64	1.53	0.54	1.00	1.92	0.39	2.14
WO × G	0.39	0.72	1.11	0.59	1.38	2.06	< 0.01	0.45	1.03	0.12	0.25	0.93
WO × G × Ant	1.09	2.11	0.65	0.25	2.98 +	1.55	0.15	3.57 +	1.54	0.53	3.93 +	1.61
WO × Hemisphere (Hem)		7.45 *	5.95 *		3.84	2.37		0.39	< 0.01		0.01	0.24
G × Hem		0.36	0.97		0.23	4.33 +		0.07	3.08 +		0.94	10.65 **
WO × Ant × Hem		3.45 *	3.37 *		0.22	0.73		1.24	0.21		0.64	0.45
G × Ant × Hem		1.77	1.91		0.14	1.19		0.28	0.57		1.12	0.74
WO × G × Hem		0.43	0.04		0.06	0.05		1.00	0.63		1.03	1.94
WO × G × Ant × Hem		0.30	1.06		0.77	0.97		1.14	2.61 +		1.56	3.30 *

+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$.

An alternative way of thinking about the SLAN's functional significance is that although it does not play an integral role in long-distance dependency formation, it is associated with the cognitive processes that are sometimes induced by long-distance dependency formation. For example, Lau (2018) has raised the possibility that the S(L)AN might reflect subvocal rehearsal performed in the phonological loop during the processing of long-distance dependency. To date, no previous ERP study has directly tested this possibility. However, there is fMRI evidence for the involvement of articulatory rehearsal in LIFG during filler-gap dependency formation. Rogalsky et al. (2008) conducted an fMRI experiment that asked their participants to perform concurrent tasks, namely articulatory tasks (e.g. whisper a consonant–vowel combination, such as "*ba da ga da*") and finger-tapping tasks (tap their fingers following certain patterns), while processing subject and object relative clauses. The results showed that the activation of the pars opercularis located in the LIFG did not increase with regard to the object relative clauses when the participants were performing the articulatory task, which impeded their articulatory rehearsal of a filler; however, it did increase during the finger tapping task, which also required a secondary task but did not involve articulatory rehearsal. However, in the present case, it is questionable whether subvocal rehearsal interacts with a discourse-level factor, leading to the SLAN's absence in a felicitous context.

A plausible possible interpretation is that this reflects the manipulation of a discourse representation in memory. As explained in the Introduction, syntactically complex sentences require certain felicity conditions to be satisfied. In OSV, the referent of the O is presupposed to be introduced by being stated either explicitly or implicitly in a preceding discourse. Thus, the processing of $O_{NEW}S_{GIVEN}V$ forced the participants to accommodate such an unsatisfied presupposition by building a coherent discourse representation, thus increasing the processing cost. By contrast, the referent of the O was explicitly introduced in the context for $O_{GIVEN}S_{NEW}V$; therefore, the participants did not need to manipulate a discourse representation.

The SLAN in an infelicitous context is reminiscent of an ERP effect that has been observed for other cases in which a presupposition is violated, namely the processing of definite NPs and pronouns without a uniquely identifiable referent (Nieuwland & Van Berkum, 2006, 2008; Van Berkum, 2004; Van Berkum et al., 1999; Van Berkum et al., 2003; Van Berkum et al., 2007; van Berkum et al., 2004). For example, Nieuwland and Van Berkum (2006) observed sustained slightly left-lateralised anterior negativity in sentences such as "*Anton*

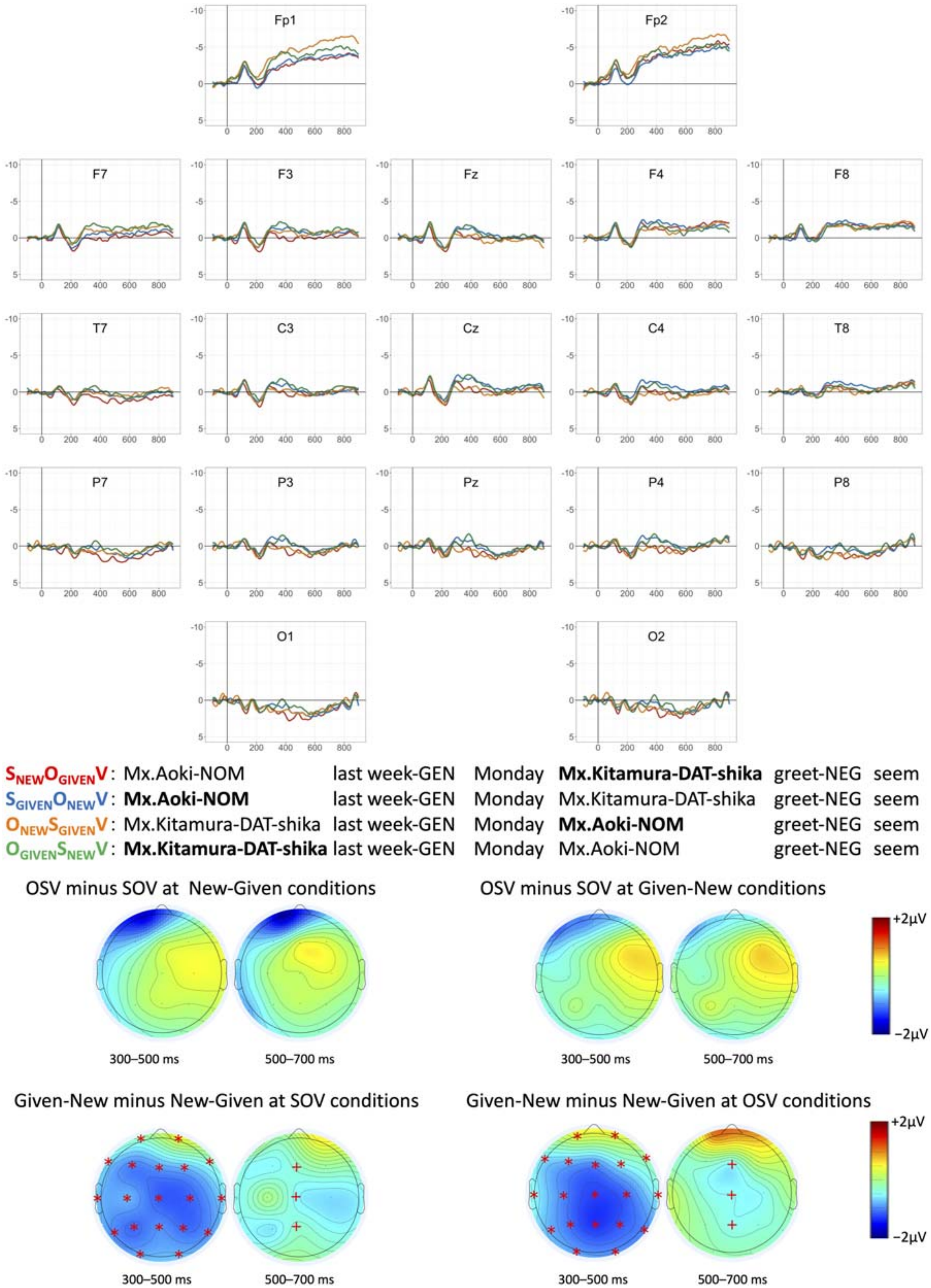

Figure 3. Grand average ERPs at NP2 (Mx. Kitamura-DAT-shika of SOV and Mx. Aoki-NOM of OSV) (Boldface in the legend indicates discourse-given NPs). The asterisks and plus marks indicate significant and marginally significant effects, respectively.

forgave Michael the problem because his car was a wreck", in which "*his*" is ambiguous in that it could refer to either Anton or Michael. Nieuwland and Van Berkum (2006) interpreted this effect as a reflection of solving an ambiguity by conducting an extensive memory search for disambiguating information. This process involves the

Table 2. Statistical results of NP2.

	300–500 ms						500–700 ms					
	Midline		Lateral		Temporal		Midline		Lateral		Temporal	
Word Order (WO)	0.03		< 0.01		2.49		0.18		0.02		1.62	
Givenness (G)	34.20	**	28.97	**	4.42	+	4.10	+	2.12		0.03	
WO × Anteriority (Ant)	0.14		0.61		2.08		0.53		0.31		1.75	
G × Ant	1.06		0.94		2.26		0.30		0.33		1.16	
WO × G	0.08		0.03		0.29		< 0.01		0.17		1.62	
WO × G × Ant	0.45		0.69		0.21		0.14		0.59		0.53	
WO × Hemisphere (Hem)			15.18	**	8.50	**			8.21	*	3.90	+
G × Hem			< 0.01		0.40				0.02		0.30	
WO × Ant × Hem			15.30	**	3.57	*			12.25	**	1.04	
G × Ant × Hem			2.16		0.85				2.45		0.84	
WO × G × Hem			< 0.01		1.34				0.04		0.75	
WO × G × Ant × Hem			0.75		0.49				2.37		0.69	

+ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$.

memory system, as the authors observed a correlation between individual working memory capacity and the magnitude of the S(L)AN effect, with a high span-group showing a larger S(L)AN.

Matchin et al. (2014) used a similar experimental paradigm in their fMRI study. They examined the processing of both backward anaphora, in which an antecedent is unavailable at the time of the pronoun's appearance ("Because he_i extinguished the flames that burned all night long, the fireman_i saved the resident"), and filler-gap dependency ("Which songs_i did the band that won the contest play $\_\_{}_i$ at the concert?"). They observed greater LIFG activation for longer distances in cases of both filler-gap and pronoun-antecedent. Although they interpreted their result as evidence that the LIFG supports prediction-based processing, the LIFG's parallel contribution to the filler-gap and to the cataphora implies that the LIFG might involve memory manipulation to build a coherent discourse model.

Obviously, it is premature to draw a conclusion with respect to SLAN's exact functional role in language comprehension. The aforementioned view needs to be tested against alternative possibilities. First, future studies need to clarify whether it can extend to the SLAN effect observed in other constructions with filler-gap dependency. The present observation of the SLAN effect in $O_{NEW}S_{GIVEN}V$ might align with the observed SLAN effect in wh-questions and relative clauses in other languages because the presuppositions in these constructions were not met in a felicitous context (Fiebach et al., 2001; King & Kutas, 1995; Müller et al., 1997; Phillips et al., 2005). In wh-questions, for example, there is an existential presupposition (Karttunen & Peters, 1976; Postal, 1971). This is evidenced by the infelicity of posing a question – for instance, "*What is Mary writing a paper about?*" – out of the blue to a hearer who does not know that there is a paper that Mary is writing. In relative clauses, which serve to restrict a referent with additional information, what is presupposed is a set of possible referents under discussion (Crain & Steedman, 1985). In the case of object relative clauses, a subject inside the relative clause should be discourse-old information that works as a restrictor for identifying a referent to whom a speaker intends to refer to (cf. Fox & Thompson, 1990). Otherwise, the use of relative clauses is not well-justified, and thus comprehenders would struggle to determine what role the relative clause plays. Roland et al. (2012) reported behavioural evidence for the context effect on relative clause processing. In contrast to the well-established observation that object relative clauses are more difficult to process than subject relative clauses (without a context), Roland et al.'s (2012) self-paced reading experiments and eye-tracking corpus analysis demonstrated that the processing disadvantage diminished or disappeared when the subject embedded in relative object clauses was a topic in prior discourse (i.e. The sculptor collected paintings. The artist that the sculptor admired …) or when it was a pronoun that is supposed to exist in discourse (i.e. The artist that I admired …). These results suggest that the discourse factor exerts an influence on relative clause processing. However, no extant ERP study on the processing of wh-questions and relative clauses examines the effect of discourse. It is therefore worth testing whether a felicitous context affects the SLAN.

Second, it is possible that the SLAN is not directly linked to language processing, instead reflecting the domain-general processes underlying language processing, such as directing attentional focus to (unexpected) information. Some memory studies in neuroscience have challenged the classic view that sustained activity reflects the short-interval retention of stimulus information (Lewis-Peacock et al., 2012; Postle, 2006). Lewis-Peacock et al. (2012) have shown that presenting a distracting stimulus (e.g. a pseudoword) while participants were keeping a task-relevant stimulus (e.g. line segments) eliminated sustained activity for the task-relevant stimulus but did not disrupt task performance (i.e. probe recognition of a stimulus feature). They argued that sustained activity only reflects information that is in attentional focus and does not serve to retain information that is outside of attentional focus. In this account, the previous results for sustained activity during a delayed interval are interpreted due to a confound of memory demand with attentional demand (i.e. the information to be remembered is also the information that is in attentional focus) (Vogel et al., 2005). In accordance with this framework, the SLAN may not play a crucial role in the successful retention of information. Instead, the input of O is likely to capture attention when it is a discourse-new NP because nothing motivates scrambling. Thus, the SLAN was observed in the non-supportive context but not in the supportive context.

4.2. Active maintenance or retrieval?

The present findings contribute to an understanding of how the language-processing system interacts with the memory system. As discussed in the Introduction, the cue-based parsing model has gained widespread acceptance in language comprehension literature (Dillon et al., 2013; Foraker & McElree, 2007, 2011; Kush et al., 2015; Lewis et al., 2006; Lewis & Vasishth, 2005; Phillips et al., 2011; Vasishth et al., 2008; Wagers et al., 2009; Wagers & McElree, 2013). In this model, the language-processing system encodes certain information associated with an NP and retrieves the NP from the content-addressable memory. The information encoded in memory is subject to decay as a function of time (Lewis et al., 2006). A controversial issue is whether filler-gap dependency is a special case of dependency formation in that the activation of a dependent element is not subject to decay (cf. Lau, 2018). Previous evidence from SLAN is taken to indicate that a dislocated element has the special status of being kept active, unlike in-situ NPs, such as a subject. Nevertheless, the present finding suggests that the SLAN reflects a cognitive process that was not specifically recruited for keeping a filler active. Only a limited volume of information is within attentional focus, whereas other information is temporarily stored in memory and needs to be retrieved by redirecting attention when it becomes necessary again. Therefore, the present finding is interpreted as evidence against the active maintenance view and consistent with the cue-based parsing models, which do not have a specialised workspace.

4.3. The absence of a P600 effect

The present experiment did not observe a P600 effect in OSV. This contrasts with Yano and Koizumi's (2018) results, in which a significant P600 effect was observed when the OSV was used infelicitously. The result is also inconsistent with previous studies on filler-gap dependency in other languages (Erdocia et al., 2009; Kaan et al., 2000; Phillips et al., 2005).

We conjecture that a possible cause of the absence of a P600 effect pertains to a quirk affecting the focus property in Japanese. It is widely assumed that an NP-shika originates within the vP domain and is licensed by a negation. Nevertheless, it has to subsequently move out of vP due to its focus property. For instance, Miyagawa et al. (2016) observed that an NP-shika cannot follow vP adverbs, such as "*umaku*" (well), as shown in (13). "*Umaku*" is assumed to be located at the edge of vP, and thus an element preceding it is supposed to be outside of vP, while an item following it should be inside vP.

(13)	Taro-wa	{*umaku}	keeki-sika	{umaku}	tsukur-anakat-ta.
	Taro-TOP	skillfully	cake-sika	skillfully	make-NEG-PST
	'Taro only made a cake well.'				

Given the theoretical claim, the focus movement of NP-shika always occurs, although such movement sometimes does not affect word order (i.e. string vacuous movement). This suggests that in SOV sentences, "*Mx.Kitamura-DAT-NPI*" undergoes a string vacuous movement, creating a short filler-gap dependency (i.e. S Adv O_i $\_\_{i}$ V). Therefore, it is possible that SOV did not provide an appropriate baseline against which to assess the P600 effect in OSV because (1) both word orders included filler-gap dependency, and (2) the integration was expected at the same timing (i.e. NP2).[10]

Alternatively, the participants might have experienced a processing overload. Considering that they had to process three proper names with no lexical bias for agent or patient, it can be supposed that the participants confused the thematic relationship and had difficulty integrating the filler O into its original position. This issue requires further investigation.

5. Conclusion

The present study used ERPs to investigate the role of discourse in filler-gap dependency formation. The result showed that OSV elicited a larger SLAN in comparison to SOV when it was used infelicitously. We interpreted this result as an indication that sustained brain activity is not necessary for processing long-distance dependency; however, it is related to manipulating a discourse model in memory when a presupposition is not met or to domain-general processes such as directing attentional focus. The present findings advance not only the understanding of SLAN's functional role, but also an understanding of how the language-processing system forms a long-distance dependency with the memory system and the way pragmatic information interacts with syntactic information during language processing.

Notes

1. The buffer is a dedicated system that stores information that will be necessary later for a short time interval.
2. The nature of stored information remains unknown, as discussed below. It may be "syntactic", since it is possible that the language-processing system not only holds a filler, but also a predicted syntactic structure that hosts a filler in working memory during the processing of filler-gap dependencies (cf. Fiebach et al., 2001; Gibson, 1998).
3. In the literature on Japanese syntax/psycholinguistics, "*shika*" is referred to as an NPI or a negative concord item (NCI). This strict distinction does not affect the present experiment; following the convention, we refer to it as an NPI.
4. Structural cases such as nominative "*-ga*" and accusative "-o" must drop when an NP is attached to "*shika*" (NP-{*ga/*o}-shika). Alternatively, non-case-marked NP-shika can be syntactically analysed as a predicate modifier. "Suzuki-shika" in the sentences (1a) and (2a) below appear to be a subject and an object, respectively, but they are not. As shown in (1b) and (2b), a nominative-marked subject and an accusative-marked object can be inserted, which suggests that "Suzuki-shika" in (1a) and (2a) are adjuncts, and the true subject and object are phonetically null (Japanese allows a pro-drop) (see Kobuchi-Philip, 2010; Shimoyama, 2011). This does not apply to case-marked NP-shika (NP-DAT-shika), which must be analysed as an argument of a verb, not an adjunct.

   (1) a. Suzuki-shika hashira-nakat-ta.
   Suzuki-shika run-NEG-PST
   'Only Suzuki ran.
   ' b. Gakusei-ga Suzuki-shika hashira-nakat-ta.
   student-NOM Suzuki-only run-NEG-PST
   'No student but only Suzuki ran.'
   (2) a. Sensei-ga Suzuki-shika shikar-anakat-ta.
   teacher-NOM Suzuki-only scold-NEG-PST
   'The teacher scolded only Suzuki.'
   b. Sensei-ga gakusei-o Suzuki-shika shikar-ana-katta.
   Teacher-NOM student-ACC Suzuki-only scold-NEG-PST
   'The teacher scolded no one, but only Suzuki.'

5. The participants were asked to indicate their preference regarding the use of their hands on the following scale: always right (100), usually right (50), both equally (0), usually left (−50), and always left (−100). Individual handedness was quantified by averaging the scores (shown in parentheses) of the ten questions.
6. The mean number of the removed ICs was 0.72. The number of the remaining trials did not differ significantly across the four conditions ($\chi^2(3) = 0.26$, $p = 0.96$; $S_{NEW}O_{GIVEN}V$: 26.7 out of 30, $S_{GIVEN}O_{NEW}V$: 26.6, $O_{NEW}S_{GIVEN}V$: 26.9, $O_{GIVEN}S_{NEW}V$: 27.4).
7. In a multi-word analysis, the baseline correction is only applied to the onset of the first word (i.e. NP1), although an epoch spans several words. The reason that the baseline correction was not applied at the onset of adverbs is that a SLAN was expected to appear at NP1 and continue until NP2. In other words, the SLAN effect was triggered by a scrambled object, and the adverbs only served as regions in which we examine whether the SLAN was indeed sustained for several words. If the baseline correction was applied to the onset of each adverb, a spurious effect would appear (cf. Steinhauer & Drury, 2012). Therefore, it is common practice to time-lock EEGs to the onset of the trigger (i.e. NP1 in the present case) and continuously examine the SLAN effect without the re-baseline correction.
8. The SLAN and P600 regions were selected based on the results of previous studies on filler-gap dependencies in Japanese. It has been demonstrated that Japanese speakers incrementally associate a filler with a gap upon encountering a nominative NP, i.e. before a verb (Aoshima et al., 2004; Hagiwara et al., 2007; Koizumi & Imamura, 2017; Ueno & Kluender, 2003). For example, Aoshima et al.'s (2004) self-paced reading experiments showed a Japanese version of the filled-gap effect prior to the input of a verb, which suggests that the filler-gap dependency is constructed at the nominative NP (Which-NP-DAT [NP-TOP [NP-NOM NP-DAT V] V]. The sentence is simplified for the purpose of illustration). This is also evident from ERP studies. For instance, Ueno and Kluender (2003) observed a P600 effect at the nominative NP of scrambled sentences (That-ACC$_i$ the reckless adventurer-NOM finally t_i discovered.). The P600 has not been observed for OSV at a verb position (Hagiwara et al., 2007; Ueno & Kluender, 2003; Yano & Koizumi, 2018). Therefore, the SLAN effect was expected to extend until the nominative NP (NP2), and the P600 was expected at the NP2, not in a verb position.
9. The simplest model was selected because more complex models that included Word Order and/or Givenness as a random slope of participant/items did not converge (final model: glmer(Cor ~ Word Order * Givenness + (1|Participant) + (1|Set), family = "binomial")).
10. This issue is not problematic for the SLAN comparison. In the SOV sentences (NP1-NOM Adv1 Adv2 NP2-DAT-shika V), the language-processing system notices the presence of "*shika*" only after encountering an NP2. Given the well-established observation that the language processing system does not construct a syntactically complex structure beyond necessity, it is plausible to assume that it processes an input string of "NP1-NOM Adv1 Adv2" (i.e. the SLAN time-windows) as a typical SOV sentence without a focus movement. Therefore, the SOV condition provides an appropriate baseline sentence for comparison with the OSV sentence in the SLAN time-window.

Acknowledgements

We are very grateful to anonymous reviewers and an editor for their insightful comments and suggestions. This study was supported by JSPS KAKENHI (#19K13182, PI: Masataka Yano, #19H05589, PI: Masatoshi Koizumi).

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by Japan Society for the Promotion of Science [grant number 19H05589,19K13182].

ORCID

Masataka Yano http://orcid.org/0000-0003-4465-8456
Masatoshi Koizumi http://orcid.org/0000-0003-0719-7128

References

Aissen, J. L. (1992). Topic and focus in Mayan. *Language*, *68*(1), 43–80. https://doi.org/10.1353/lan.1992.0017

Aoshima, S., Phillips, C., & Weinberg, A. (2003). Processing of Japanese WH-scrambling constructions. In W. McClure (Ed.), *Japanese/Korean linguistics 12* (Vol. 12, pp. 179–191). Stanford, CA: CSLI Publications.

Aoshima, S., Phillips, C., & Weinberg, A. (2004). Processing filler-gap dependencies in a head-final language. *Journal of Memory and Language*, *51*(1), 23–54. https://doi.org/10.1016/j.jml.2004.03.001

Baddeley, A. D., & Hitch, G. (1974). Working memory. *Psychology of Learning and Motivation*, *8*, 47–89. https://doi.org/10.1016/S0079-7421(08)60452-1

Batuev, A. S., Shaefer, V. I., & Orlov, A. A. (1985). Comparative characteristics of unit activity in the prefrontal and parietal areas during delayed performance in monkeys. *Behavioural Brain Research*, *16*(1), 57–70. https://doi.org/10.1016/0166-4328(85)90082-8

Birner, B. J., & Ward, G. (2009). Information structure and syntactic structure. *Language and Linguistics Compass*, *3*(4), 1167–1187. https://doi.org/10.1111/j.1749-818X.2009.00146.x

Burkhardt, P. (2006). Inferential bridging relations reveal distinct neural mechanisms: Evidence from event-related brain potentials. *Brain and Language*, *98*(2), 159–168. https://doi.org/10.1016/j.bandl.2006.04.005

Burkhardt, P. (2007). The P600 refects cost of new information in discourse memory. *NeuroReport*, *18*(17), 1851–1854. https://doi.org/10.1097/WNR.0b013e3282f1a999

Clahsen, H., & Featherston, S. (1999). Antecedent priming at trace positions: Evidence from German scrambling. *Journal of Psycholinguistic Research*, *28*(4), 415–437. https://doi.org/10.1023/A:1023293132656

Clifton, C., & Frazier, L. (2004). Should given information come before new? Yes and no. *Memory and Cognition*, *32*(6), 886–895. https://doi.org/10.3758/BF03196867

Crain, S., & Fodor, J. D. (1985). How can grammars help parsers? In D. R. Dowty, L. Karttunen, & A. M. Zwicky (Eds.), *Natural language parsing psychological, computational, and theoretical perspectives* (pp. 94–128). Cambridge University Press.

Crain, S., & Steedman, M. (1985). On not being led up the garden path: The use of context by the psychological syntax processor. In D. R. Dowty, L. Karttunen, & A. M. Zwicky (Eds.), *Natural language parsing* (pp. 320–358). Cambridge University Press.

Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, *134*(1), 9–21. https://doi.org/10.1016/j.jneumeth.2003.10.009

Dillon, B., Mishler, A., Sloggett, S., & Phillips, C. (2013). Contrasting intrusion profiles for agreement and anaphora: Experimental and modeling evidence. *Journal of Memory and Language*, *69*(2), 85–103. https://doi.org/10.1016/j.jml.2013.04.003

Erdocia, K., Laka, I., Mestres-Missé, A., & Rodriguez-Fornells, A. (2009). Syntactic complexity and ambiguity resolution in a free word order language: Behavioral and electrophysiological evidences from Basque. *Brain and Language*, *109*(1), 1–17. https://doi.org/10.1016/j.bandl.2008.12.003

Ferreira, V. S., & Yoshita, H. (2003). Given-new ordering effects on the production of scrambled sentences in Japanese. *Journal of Psycholinguistic Research*, *32*(6), 669–692. https://doi.org/10.1023/A:1026146332132

Fiebach, C. J., Schlesewsky, M., & Friederici, A. D. (2001). Syntactic working memory and the establishment of filler-gap dependencies: Insights from ERPs and fMRI. *Journal of Psycholinguistic Research*, *30*(3), 321–338. https://doi.org/10.1023/A:1010447102554

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*(2), 250–272. https://doi.org/10.1016/S0749-596X(02)00004-9

Fiebach, C. J., Schlesewsky, M., Lohmann, G., Von Cramon, D. Y., & Friederici, A. D. (2005). Revisiting the role of Broca's area in sentence processing: Syntactic integration versus syntactic working memory. *Human Brain Mapping*, *24*(2), 79–91. https://doi.org/10.1002/hbm.20070

Foraker, S., & McElree, B. (2007). The role of prominence in pronoun resolution: Active versus passive representations. *Journal of Memory and Language*, *56*(3), 357–383. https://doi.org/10.1016/j.jml.2006.07.004

Foraker, S., & McElree, B. (2011). Comprehension of linguistic dependencies: Speed-accuracy tradeoff evidence for direct-access retrieval from memory. *Language and Linguistics Compass*, *5*(11), 764–783. https://doi.org/10.1111/j.1749-818X.2011.00313.x

Fox, B. A., & Thompson, S. A. (1990). A discourse explanation of the grammar of relative clauses in English conversation. *Language*, *66*(2), 297–316. https://doi.org/10.2307/414888

Frazier, L. (1987). Syntactic processing: Evidence from Dutch. *Natural Language and Linguistic Theory*, *5*(4), 519–559. https://doi.org/10.1007/BF00138988

Frazier, L., & Clifton, C. (1989). Successive cyclicity in the grammar and the parser. *Language and Cognitive Processes*, *4*(2), 93–126. https://doi.org/10.1080/01690968908406359

Frazier, L., & Flores d'Arcais, G. B. (1989). Filler driven parsing: A study of gap filling in Dutch. *Journal of Memory and Language*, *28*(3), 331–344. https://doi.org/10.1016/0749-596X(89)90037-5

Funahashi, S., Bruce, C. J., & Goldman-Rakic, P. S. (1989). Mnemonic coding of visual space in the monkey's dorsolateral prefrontal cortex. *Journal of Neurophysiology*, *61*(2), 331–349. https://doi.org/10.1152/jn.1989.61.2.331

Gibson, E. (1998). Linguistic complexity: Locality of syntactic dependencies. *Cognition*, *68*(1), 1–76. https://doi.org/10.1016/S0010-0277(98)00034-1

Greenhouse, S. W., & Geisser, S. (1959). On methods in the analysis of profile data. *Psychometrika*, *24*(2), 95–112. https://doi.org/10.1007/BF02289823

Grodner, D., Gibson, E., & Watson, D. (2005). The influence of contextual contrast on syntactic processing: Evidence for strong-interaction in sentence comprehension. *Cognition*, *95*(3), 275–296. https://doi.org/10.1016/j.cognition.2004.01.007

Grodzinsky, Y., & Santi, A. (2008). The battle for Broca's region. *Trends in Cognitive Sciences*, *12*(12), 474–480. https://doi.org/10.1016/j.tics.2008.09.001

Hagiwara, H., Soshi, T., Ishihara, M., & Imanaka, K. (2007). A topographical study on the event-related potential correlates of scrambled word order in Japanese complex sentences. *Journal of Cognitive Neuroscience*, *19*(2), 175–193. https://doi.org/10.1162/jocn.2007.19.2.175

Heim, I. (1982). *The semantics of definite and indefinite NPs*. University of Massachusetts at Amherst dissertation.

Imamura, S. (2014). The influence of givenness and heaviness on OSV in Japanese. *Proceedings of the 28th Pacific Asia Conference on Language, Information and Computation*, pp. 224–233.

Imamura, S. (2015). The effects of givenness and heaviness on VP-internal scrambling and VP-external scrambling in Japanese. *Studies in Pragmatics*, *17*, 1–16.

Imamura, S., & Koizumi, M. (2011). A centering analysis of word order in Japanese. *Tohoku Studies in Linguistics*, *20*, 59–74.

Imamura, S., Sato, Y., & Koizumi, M. (2016). The processing cost of scrambling and topicalization in Japanese. *Frontiers in Psychology*, *7*. https://doi.org/10.3389/fpsyg.2016.00531

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*(2), 159–201. https://doi.org/10.1080/016909600386084

Kaiser, E., & Trueswell, J. C. (2004). The role of discourse context in the processing of a flexible word-order language. *Cognition*, *94*(2), 113–147. https://doi.org/10.1016/j.cognition.2004.01.002

Karttunen, L., & Peters, S. (1976). What indirect questions conventionally implicate. *CLS 12: Papers from the Twelfth Regional Meeting*, pp. 351–368. Chicago Linguistic Society.

Kataoka, K. (2010). Hitei kyokusei-to tougo zyoken [Negative polarity and sytactic condition]. In Y. Kato, A. Yoshimura, & I. Imani (Eds.), *Hitei-to gengo riron* [Negation and linguistic theory] (pp. 118–140). Kaitakusha.

Kato, Y. (1985). *Negative sentences in Japanese*. Sophia University.

Kato, Y. (1994). Negative polarity and movement. *MIT Working Papers in Linguistics*, *24*, 101–120.

Kim, J., Koizumi, M., Ikuta, N., Fukumitsu, Y., Kimura, N., Iwata, K., Fukumitsu, Y., Kimura, N., Iwata, K., Watanabe, J., Yokoyama, S., Sato, S., Horie, K., & Kawashima, R. (2009). Scrambling effects on the processing of Japanese sentences: An fMRI study. *Journal of Neurolinguistics*, *22*(2), 151–166. https://doi.org/10.1016/j.jneuroling.2008.07.005

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*, *7*(3), 376–395. https://doi.org/10.1162/jocn.1995.7.3.376

Kinno, R., Kawamura, M., Shioda, S., & Sakai, K. L. (2008). Neural correlates of noncanonical syntactic processing revealed by a picture-sentence matching task. *Human Brain Mapping*, *29*(9), 1015–1027. https://doi.org/10.1002/hbm.20441

Kluender, R., & Kutas, M. (1993a). Bridging the gap: Evidence from ERPs on the processing of unbounded dependencies. *Journal of Cognitive Neuroscience*, *5*(2), 196–214. https://doi.org/10.1162/jocn.1993.5.2.196

Kluender, R., & Kutas, M. (1993b). Subjacency as a processing phenomenon. *Language and Cognitive Processes*, *8*(4), 573–633. https://doi.org/10.1080/01690969308407588

Kobuchi-Philip, M. (2010). Japanese mo \`also/even' and shika \`except for/only'. *Proceedings of Sinn Und Bedeutung*, *14*, 219–236.

Koizumi, M., & Imamura, S. (2017). Interaction between syntactic structure and information structure in the processing of a head-final language. *Journal of Psycholinguistic Research*, *46*(1), 247–260. https://doi.org/10.1007/s10936-016-9433-3

Kuno, S. (1987). *Functional syntax: Anaphora, discourse, and empathy*. University of Chicago Press.

Kush, D., Lidz, J., & Phillips, C. (2015). Relation-sensitive retrieval: Evidence from bound variable pronouns. *Journal of Memory and Language*, *82*, 18–40. https://doi.org/10.1016/j.jml.2015.02.003

Kutas, M., & Federmeier, K. D. (2000). Electrophysiology reveals semantic memory use in language comprehension. *Trends in Cognitive Sciences*, *4*(12), 463–470. https://doi.org/10.1016/S1364-6613(00)01560-6

Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: Finding meaning in the N400 component of the event-related brain potential (ERP). *Annual Review of Psychology*, *62*(1), 621–647. https://doi.org/10.1146/annurev.psych.093008.131123

Lau, E. F. (2018). Neural indices of structured sentence representation: State of the art. *Psychology of Learning and Motivation*, *68*, 117–142. https://doi.org/10.1016/bs.plm.2018.08.004

Lau, E. F., Almeida, D., Hines, P. C., & Poeppel, D. (2009). A lexical basis for N400 context effects: Evidence from MEG. *Brain and Language*, *111*(3), 161–172. https://doi.org/10.1016/j.bandl.2009.08.007

Lau, E. F., & Liao, C. H. (2018). Linguistic structure across time: ERP responses to coordinated and uncoordinated noun phrases. *Language, Cognition and Neuroscience*, *33*(5), 633–647. https://doi.org/10.1080/23273798.2017.1400081

Lau, E. F., Phillips, C., & Poeppel, D. (2008). A cortical network for semantics: (De)constructing the N400. *Nature Reviews Neuroscience*, *9*(12), 920–933. https://doi.org/10.1038/nrn2532

Lee, M. W. (2004). Another look at the role of empty categories in sentence processing (and grammar). *Journal of*

Psycholinguistic Research, *33*(1), 51–73. https://doi.org/10.1023/B:JOPR.0000010514.50468.30

Lewis-Peacock, J. A., Drysdale, A. T., Oberauer, K., & Postle, B. R. (2012). Neural evidence for a distinction between short-term memory and the focus of attention. *Journal of Cognitive Neuroscience*, *24*(1), 61–79. https://doi.org/10.1162/jocn_a_00140

Lewis, R. L., & Vasishth, S. (2005). An activation-based model of sentence processing as skilled memory retrieval. *Cognitive Science*, *29*(3), 375–419. https://doi.org/10.1207/s15516709cog0000_25

Lewis, R. L., Vasishth, S., & Van Dyke, J. A. (2006). Computational principles of working memory in sentence comprehension. *Trends in Cognitive Sciences*, *10*(10), 447–454. https://doi.org/10.1016/j.tics.2006.08.007

Lopez-Calderon, J., & Luck, S. J. (2014). ERPLAB: An open-source toolbox for the analysis of event-related potentials. *Frontiers in Human Neuroscience*, *8*, 213. https://doi.org/10.3389/fnhum.2014.00213

Matchin, W., Sprouse, J., & Hickok, G. (2014). A structural distance effect for backward anaphora in Broca's area: An fMRI study. *Brain and Language*, *138*, 1–11. https://doi.org/10.1016/j.bandl.2014.09.001

Matzke, M., Mai, H., Nager, W., Rüsseler, J., & Münte, T. F. (2002). The costs of freedom: An ERP – Study of non-canonical sentences. *Clinical Neurophysiology*, *113*(6), 844–852. https://doi.org/10.1016/S1388-2457(02)00059-7

Meng, M., Bader, M., & Bayer, J. (1999). Die Verarbeitung von Subjekt– Objekt Ambiguitäten im Kontext [The processing of subject-object ambiguities in context]. *Proceedings Der 4. Fachtagung Der Gesellschaft Für Kognitionswissenschaft*, pp. 244–249.

Miyagawa, S., Nishioka, N., & Zeijlstra, H. (2016). Negative sensitive items and the discourse-configurational nature of Japanese. *Glossa: A Journal of General Linguistics*, *1*(1), 33. https://doi.org/10.5334/gjgl.6

Miyamoto, E., & Takahashi, S. (2002). Antecedent reactivation in the processing of scrambling in Japanese. *MIT Working Papers in Linguistics*, *43*, 123–138.

Müller, H. M., King, J. W., & Kutas, M. (1997). Event-related potentials elicited by spoken relative clauses. *Cognitive Brain Research*, *5*(3), 193–203. https://doi.org/10.1016/S0926-6410(96)00070-5

Nakano, Y., Felser, C., & Clahsen, H. (2002). Antecedent priming at trace positions in Japanese long-distance scrambling. *Journal of Psycholinguistic Research*, *31*(5), 531–571. https://doi.org/10.1023/A:1021260920232

Nicol, J., & Swinney, D. (1989). The role of structure in coreference assignment during sentence comprehension. *Journal of Psycholinguistic Research*, *18*(1), 5–19. https://doi.org/10.1007/BF01069043

Nieuwland, M. S., & Van Berkum, J. J. A. (2006). Individual differences and contextual bias in pronoun resolution: Evidence from ERPs. *Brain Research*, *1118*(1), 155–167. https://doi.org/10.1016/j.brainres.2006.08.022

Nieuwland, M. S., & Van Berkum, J. J. A. (2008). The interplay between semantic and referential aspects of anaphoric noun phrase resolution: Evidence from ERPs. *Brain and Language*, *106*(2), 119–131. https://doi.org/10.1016/j.bandl.2008.05.001

Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*, *9*(1), 97–113. https://doi.org/10.1016/0028-3932(71)90067-4

Omaki, A., Lau, E. F., Davidson White, I., Dakan, M. L., Apple, A., & Phillips, C. (2015). Hyper-active gap filling. *Frontiers in Psychology*, *06*, 384. https://doi.org/10.3389/fpsyg.2015.00384

Öztekin, I., Davachi, L., & McElree, B. (2010). Are representations in working memory distinct from representations in long-term memory? Neural evidence in support of a single Store. *Psychological Science*, *21*(8), 1123–1133. https://doi.org/10.1177/0956797610376651

Palmer, J. A., & Makeig, S. (2008). Newton method for the ICA mixture model. *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, pp. 1805–1808.

Phillips, C. (2006). The real-time status of island phenomena. *Language*, *82*(4), 795–823. https://doi.org/10.1353/lan.2006.0217

Phillips, C., Kazanina, N., & Abada, S. H. (2005). ERP effects of the processing of syntactic long-distance dependencies. *Cognitive Brain Research*, *22*(3), 407–428. https://doi.org/10.1016/j.cogbrainres.2004.09.012

Phillips, C., Wagers, M. W., & Lau, E. F. (2011). Grammatical illusions and selective fallibility in real-time language comprehension. In J. Runner (Ed.), *Experiments at the interfaces* (Vol. 37, pp. 147–180). Brill.

Pion-Tonachini, L., Kreutz-Delgado, K., & Makeig, S. (2019). ICLabel: An automated electroencephalographic independent component classifier, dataset, and website. *NeuroImage*, *198*, 181–197. https://doi.org/10.1016/j.neuroimage.2019.05.026

Postal, P. M. (1971). *Cross-over phenomena*. Rinehart and Winston.

Postle, B. R. (2006). Working memory as an emergent property of the mind and brain. *Neuroscience*, *139*(1), 23–38. https://doi.org/10.1016/j.neuroscience.2005.06.005

Rogalsky, C., Almeida, D., Sprouse, J., & Hickok, G. (2015). Sentence processing selectivity in Broca's area: Evident for structure but not syntactic movement*, Language, Cognition and Neuroscience*, *30*(10), 1326–1338. https://doi.org/10.1080/23273798.2015.1066831

Rogalsky, C., Matchin, W., & Hickok, G. (2008). Broca's area, sentence comprehension, and working memory: An fMRI study. *Frontiers in Human Neuroscience*, *2*, 14. https://doi.org/10.3389/neuro.09.014.2008

Roland, D., Mauner, G., Meara, C. O., & Yun, H. (2012). Discourse expectations and relative clause processing. *Journal of Memory and Language*, *66*(3), 479–508. https://doi.org/10.1016/j.jml.2011.12.004

Saito, M. (1985). *Some asymmetries in Japanese and their theoretical implications*. [Doctoral dissertation]. MIT.

Santi, A., & Grodzinsky, Y. (2007). Working memory and syntax interact in Broca's area. *NeuroImage*, *37*(1), 8–17. https://doi.org/10.1016/j.neuroimage.2007.04.047

Santi, A., & Grodzinsky, Y. (2010). fMRI adaptation dissociates syntactic complexity dimensions. *NeuroImage*, *51*(4), 1285–1293. https://doi.org/10.1016/j.neuroimage.2010.03.034

Sekerina, I. A. (2003). Scrambling and processing: Dependencies, complexity and constraints. In S. Karimi (Ed.), *Word order and scrambling* (pp. 301–324). Blackwell.

Shimoyama, J. (2011). Japanese indeterminate negative polarity items and their scope. *Journal of Semantics*, *28*(4), 413–450. https://doi.org/10.1093/jos/ffr004

Steinhauer, K., & Drury, J. E. (2012). On the early left-anterior negativity (ELAN) in syntax studies. *Brain and Language*, *120*(2), 135–162. https://doi.org/10.1016/j.bandl.2011.07.001

Stowe, L. A. (1986). Parsing wh-constructions: Evidence for on-line gap location. *Language and Cognitive Processes*, *1*(3), 227–245. https://doi.org/10.1080/01690968608407062

Swinney, D., Prather, P., & Love, T. (2000). The time-course of lexical access and the role of context: Converging evidence from normal and aphasic processing. In Y. Grodzinsky, L. Shapiro, & D. A. Swinney (Eds.), *Language and the brain*: Representation and processing (pp. 273–290). Academic Press.

Ueno, M., & Kluender, R. (2003). Event-related brain indices of Japanese scrambling. *Brain and Language*, *86*(2), 243–271. https://doi.org/10.1016/S0093-934X(02)00543-6

Ueno, M., & Kluender, R. (2009). On the processing of Japanese wh-questions: An ERP study. *Brain Research*, *1290*, 63–90. https://doi.org/10.1016/j.brainres.2009.05.084

Van Berkum, J. J. A. (2004). Sentence comprehension in a wider discourse: Can we use ERPs to keep track of things? In M. Carreiras & C. J. Clifton (Eds.), *The on-line study of sentence comprehension: Eyetracking, ERPs and beyond* (pp. 229–270). Psychology Press.

Van Berkum, J. J. A., Brown, C. M., & Hagoort, P. (1999). Early referential context effects in sentence processing: Evidence from event-related brain potentials. *Journal of Memory and Language*, *41*(2), 147–182. https://doi.org/10.1006/jmla.1999.2641

Van Berkum, J. J. A., Brown, C. M., Hagoort, P., & Zwitserlood, P. (2003). Event-related brain potentials reflect discourse-referential ambiguity in spoken language comprehension. *Psychophysiology*, *40*(2), 235–248. https://doi.org/10.1111/1469-8986.00025

Van Berkum, J. J. A., Koornneef, A. W., Otten, M., & Nieuwland, M. S. (2007). Establishing reference in language comprehension: An electrophysiological perspective. *Brain Research*, *1146*, 158–171. https://doi.org/10.1016/j.brainres.2006.06.091

van Berkum, J. J. A., Zwitserlood, P., Bastiaansen, M. C. M., Brown, C. M., & Hagoort, P. (2004). So who's "he" anyway? Differential ERP and ERSP effects of referential success, ambiguity and failure during spoken language comprehension. *Annual Meeting of the Cognitive Neuroscience Society*, *16*, 70. San Francisco, April 18–20.

Van Dyke, J. A., & McElree, B. (2006). Retrieval interference in sentence comprehension. *Journal of Memory and Language*, *55*(2), 157–166. https://doi.org/10.1016/j.jml.2006.03.007

Vasishth, S., Brussow, S., Lewis, R., & Drenhaus, H. (2008). Processing polarity: How the ungrammatical intrudes on the grammatical. *Cognitive Science*, *32*(4), 685–712. https://doi.org/10.1080/03640210802066865

Vogel, E. K., McCollough, A. W., & Machizawa, M. G. (2005). Neural measures reveal individual differences in controlling access to working memory. *Nature*, *438*(7067), 500–503. https://doi.org/10.1038/nature04171

Wagers, M. W., Lau, E. F., & Phillips, C. (2009). Agreement attraction in comprehension: Representations and processes. *Journal of Memory and Language*, *61*(2), 206–237. https://doi.org/10.1016/j.jml.2009.04.002

Wagers, M. W., & McElree, B. (2013). Working memory and language processing: Theory, data, and directions for future research. In C. Boeckx & K. K. Grohmann (Eds.), *The Cambridge Handbook of Biolinguistics* (pp. 203–232). Cambridge University Press.

Wagers, M. W., & Phillips, C. (2014). Going the distance: Memory and control processes in active dependency construction. *Quarterly Journal of Experimental Psychology*, *67*(7), 1274–1304. https://doi.org/10.1080/17470218.2013.858363

Wanner, H., & Maratsos, M. (1978). An ATN approach to comprehension. In M. Halle, J. Bresnan, & G. A. Miller (Eds.), *Linguistic theory and psychological reality* (pp. 119–161). MIT Press.

Yano, M. (2019). On the nature of the discourse effect on extraction in Japanese. *Glossa: A Journal of General Linguistics*, *4*(1)), https://doi.org/10.5334/gjgl.822

Yano, M., & Koizumi, M. (2018). Processing of non-canonical word orders in (in)felicitous contexts: Evidence from event-related brain potentials. *Language, Cognition and Neuroscience*, *33*(10), 1340–1354. https://doi.org/10.1080/23273798.2018.1489066