Published in final edited form as:

Brain Res. 2009 September 22; 1290: 63–90. doi:10.1016/j.brainres.2009.05.084.

On the Processing of Japanese Wh-Questions: An ERP Study

Mieko Ueno and

Department of Linguistics, University of California, San Diego

Robert Kluender

Department of Linguistics, University of California, San Diego

Abstract

The processing of Japanese wh-questions was investigated using event-related brain potentials (ERPs). Unlike in English or German, a wh-element in Japanese need not be displaced from its canonical position, but instead needs a corresponding Q(uestion)-particle to indicate its interrogative scope. We tested to see if there were any processing correlates specific to these features of Japanese wh-questions. Both mono-clausal and bi-clausal Japanese wh-questions elicited right-lateralized anterior negativity (RAN) between wh-words and corresponding Q-particles, relative to structurally-equivalent yes/no-question control conditions. These results suggest a reliable neural processing correlate of the dependency between wh-elements and Q-particles in Japanese, which is similar to effects of (left) anterior negativity seen between wh-fillers and gaps in English and German, but with a right- rather than left-lateralized distribution. It is suggested that wh-in-situ questions in Japanese are processed by the incremental formation of a long-distance dependency between wh-elements and their Q-particles, resulting in a working memory load for keeping track of the scopeless wh-elements.

Keywords

ERP; LAN; RAN; filler-gap dependencies; Japanese sentence processing; P200

1. Introduction

1.1 Wh-movement languages and filler-gap dependencies

Wh-questions—i.e., questions that contain wh-elements such as ‘what’ and ‘who’—have long been a focus of linguistic research not only in the field of theoretical syntax (Chomsky, 1981, 1986, 1995, among many others) but also in the field of language processing (e.g., Fodor, 1978, 1989; Crain and Fodor, 1985; Stowe, 1986; Frazier and Clifton, 1989; de Vincenzi, 1991, among others). In terms of electrophysiological processes, there have been a number of studies that have investigated the processing of wh-questions using ERPs (Kluender and Kutas, 1993a, b; McKinnon and Osterhout, 1996; Müller et al., 1997; Kluender and Münte, 1998; Kaan et al., 2000; Fiebach et al., 2001, 2002; Felser et al., 2003; Phillips et al., 2005; Gouvea et al., submitted). These studies have been done in English and German, both of which are so-

Address correspondence to: Mieko Ueno, Department of Linguistics, University of California, San Diego, 9500 Gilman Drive #108, La Jolla, CA 92093-0108, phone: (217) 721-8227, fax: (858) 534-4789, ueno@ling.ucsd.edu.

The experiments were undertaken with the understanding and written consent of each subject.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

called “wh-movement” languages that require wh-elements to be displaced to the beginning of a clause.

Consider (1a): the wh-object *what* in English must be displaced to the beginning of its clause (except in the case of an “echo question” like *Calvin brought WHAT?*), rather than appear in canonical post-verbal position, as does the non-wh object *pizza* in (1b).

(1)

a. Wh-question

What did Calvin bring __ ?

FILLER GAP

b. Yes/no-question

Did Calvin bring pizza?

In the psycholinguistic literature, the displaced wh-element is called a “filler” while the canonical position is called a “gap”, and these are said to be dependent on each other for successful sentence interpretation. This is said to be because a filler is generally ambiguous in terms of its grammatical function (such as “subject” or “object”) and semantic role (such as “agent” or “patient”) until the parser hits the gap position and is able to unambiguously identify this missing information (cf. Fodor, 1989). The processing of this “filler-gap dependency” has been associated with an ERP component known as left anterior negativity (LAN). The LAN is a negative voltage deflection that is larger over the front of the head than over the back and is often left-lateralized. It has been observed in a transient/phasic form, with a duration of a few hundred milliseconds, and in a sustained slow wave/potential form, with a duration of several seconds. Besides phrase structure and morphosyntactic violations (e.g., Neville et al., 1991; Coulson et al., 1998; Martín-Loeches et al., 2005; see Vos et al., 2001 for a review), LAN has been linked with increased working memory load in associating a displaced wh-filler with its gap (e.g., Kluender and Kutas, 1993a, b; King and Kutas, 1995a; Müller et al., 1997; Kluender and Münte, 1998). For instance, Kluender and Kutas (1993a, b) reported phasic LAN effects at the positions immediately following the filler and the gap of various English object wh-questions, and concluded that both the storage of a filler in working memory and its retrieval¹ for filler-gap assignment are indexed by LAN effects. Likewise, King and Kutas (1995a) reported a relatively frontal, negative slow wave between the filler and the gap, as well as a phasic LAN effect immediately following the gap, in response to English object relative clause sentences.

More recent studies have reported P600 effects instead of LAN effects at the gap location. The P600 is a positivity that typically peaks 500–800 ms after stimulus onset and is typically broadly distributed over the head with a bilateral centro-posterior maximum. For instance, Kaan et al. (2000) found P600 effects at the pre-gap position of wh-questions (*bring* in 1a) in comparison to yes/no-questions. Based on this, they argued that the P600 indexes the difficulty of syntactically integrating wh-fillers in the on-going parse (see also Gouvea et al., submitted). Other recent studies have reported the combination of both slow/phasic LAN and P600 effects in wh-questions in English (Phillips et al., 2005) and German (Fiebach et al., 2001; Fiebach et al., 2002; Felser et al., 2003), as well as in “scrambled” (see Section 1.2 below) wh-questions in Japanese (Ueno and Kluender, 2003).

¹This process has been recently argued to be the parser’s backward search through memory for an appropriate entity to establish a dependency (Kwon, 2008 based on data for Korean relative clauses; see Ueno and Garnsey, 2008 for a similar LAN effect for Japanese relative clauses).

1.2 Japanese: A wh-in-situ language

Unlike English and German, Japanese is a “wh-in-situ” language in which wh-words stay in the same canonical SOV (subject-object-verb) position as their non-wh counterparts. As shown in (2), ‘pizza’ and ‘what’ typically occupy the same position.

(2)

a.	カルビンが
Calvin-ga
Calvin-NOM(INATIVE) | ピザを
pizza-o
pizza-ACC(USATIVE) | 持ってきたんですか。
mottekita-ndesu-ka.
brought-POLITE-Q(UESTION) |
|----|---|--------------------------------------|--|

‘Did Calvin bring pizza?’

b.	カルビンが
Calvin-ga
Calvin-NOM | 何を
nani-o
what-ACC | 持ってきたんですか。
mottekita-ndesu-ka.
brought-POLITE-Q |
|----|----------------------------------|--|---|

‘What did Calvin bring?’

Although it is also possible in Japanese to displace objects to the beginning of the clause, this involves another process called “scrambling” (see Saito 1985, 1992 for syntactic considerations and Yamashita 1997, 2002, Ueno and Kluender, 2003, and Hagiwara et al., 2007 for processing considerations).

While displacement/scrambling is optional, Japanese wh-words always require a question (Q) particle *ka* or *no* (meaning ‘whether’) at the end of the clause, as in (3a). The only exception is when a question is spoken with a rising intonation, which can be interpreted as a prosodic version of a Q-particle. Wh-questions without a Q-particle are ungrammatical in Japanese, as shown in (3b). (Here a non-question particle *-yo* ‘you.know’ is inserted after the verb to block possible rising intonation imposed implicitly by the reader.)

(3)

a.	カルビンが
Calvin-ga
Calvin-NOM | 何を
nani-o
what-ACC | 持ってきたんですか。
mottekita-ndesu-ka.
brought-POLITE-Q |
|----|----------------------------------|--|---|

‘What did Calvin bring?’

b.	カルビンが
*Calvin-ga
Calvin-NOM | 何を
nani-o
what-ACC | 持ってきたんですよ。
mottekita-ndesu-yo.
brought-POLITE-you.know |
|----|-----------------------------------|--|--|

*‘Calvin brought what.’

In addition, the Q-particle *ka* determines the interrogative scope of a wh-element. Interrogative scope can be defined as the domain of the sentence that is being questioned. In wh-movement languages, such as English and German, the position of a wh-element at the beginning of a clause transparently indicates its interrogative scope within the sentence. For instance, both (4a) and (4b) consist of two clauses, main and embedded. The wh-element *what* can be placed either at the beginning of the embedded clause as in (4a), yielding an embedded clause wh-question (traditionally termed an “indirect question”)², or at the beginning of the main clause as in (4b), yielding a main clause wh-question (also termed a “direct question”). Although in daily conversation one might actually provide the referent of the wh-word (*pizza*) in answer to either type of question for pragmatic reasons, the logical response to the two is different. The

logical answer to an embedded clause wh-question like (4a) would be *yes* or *no*, since no element in the main clause is questioned by a wh-word, while the logical answer to a matrix clause wh-question like (4b) would be the referent of the wh-word (*pizza* in this case), and cannot be *yes* or *no*.

(4) Wh-movement languages (e.g., English)

a. Embedded clause wh-question

[Did Hobbes ask [what Calvin brought ___]]? Logical Answer: Yes, he did.

b. Main clause wh-question

[What did Hobbes say [Calvin brought ___]]? Logical Answer: Pizza.

In Japanese, the position of the Q-particle rather than the position of the wh-element indicates the interrogative scope. While the wh-elements in the indirect (embedded clause) question (5a) and the direct (main clause) question (5b) both remain in the embedded clause, the Q-particle *ka* appears at the end of the clause bearing interrogative scope, either the embedded (5a) or the main clause (5b), yielding an embedded (5a) or main clause (5b) wh-question.

(5)

a. Embedded clause wh-question

				embedded-Q	
				↓	
ホッブスは	カルビンが	何を	持ってきたか		聞きましたか。
[Hobbes-wa	[Calvin-ga	nani-o	mottekita-ka]		kiki-mashi-ta-ka]
Hobbes-TOPIC	Calvin-NOM	what-ACC	brought-Q		ask-POLITE-PAST-Q

‘Did Hobbes ask what Calvin brought?’

b. Main clause wh-question

					main-Q
					↓
ホッブスは	カルビンが	何を	持ってきたと		言いましたか。
[Hobbes-wa	[Calvin-ga	nani-o	mottekita-to]		ii-mashi-ta-ka]
Hobbes-TOPIC	Calvin-NOM	what-ACC	brought-that		say-POLITE-PAST-Q

‘What did Hobbes say Calvin brought?’

For main clause wh-questions like (5b), a non-question particle *to* ‘that’ is attached to the embedded verb.

Thus, in English a wh-element is displaced to the left of its gap and marks its own scope as shown in (6a), while in Japanese a Q-particle appears to the right of the corresponding wh-element to mark its scope (6b).

(6)

a. English

²More strictly speaking, (4a) is both an embedded clause wh-question and a (main clause) yes/no-question. The embedded clause wh-question only would be *Hobbes asked what Calvin brought*. However, the form shown in (4a) is used because it is more parallel to the ERP stimuli discussed later.

scope marker

[**What** did Hobbes say [Calvin brought]]?
 FILLER GAP

b. Japanese

scope marker

[Hobbes-wa [Calvin-ga **nani-o** mottekita-to] ii-mashi-ta-ka]?
 Hobbes-TOPIC Calvin-NOM what-ACC brought-that say-POLITE-PAST-Q

‘What did Hobbes say Calvin brought?’

In this sense, the wh-scope marking system in Japanese can be said to be almost a mirror image of the wh-scope marking system in English.

1.3 Processing of Japanese wh-questions

As discussed above, Japanese wh-elements typically stay in their canonical subject-object-verb positions, and in addition, case-markers (e.g., *-ga* ‘NOMINATIVE’, *-o* ‘ACCUSATIVE’) are usually attached to wh- (as well as non-wh-) nouns. Therefore, the semantic role and the grammatical function of a Japanese wh-element are typically unambiguous. However, what is ambiguous is its interrogative scope. Unlike the processing of wh-questions in wh-movement languages like English or German, where the parser can immediately read the interrogative scope relations off the position of wh-elements, the parser in a wh-in-situ language like Japanese is unable to determine the scope of a wh-element until the related Q-particle is encountered.

Perhaps because of this necessity, experimental evidence suggests a close relationship between Japanese wh-elements and Q-particles. In an ERP study, Nakagome et al. (2001) examined the processing of mono-clausal wh-questions with (grammatical) and without (ungrammatical) Q-particles at the sentence-final verb position, and reported a P600 effect in response to wh-questions without Q-particles. In a self-paced reading experiment, Miyamoto and Takahashi (2002) reported that main clause wh-questions like (5b) were read more slowly than embedded clause wh-questions (like 5a) at the embedded verb position, i.e., ‘verb-that’ without a Q-particle was read more slowly than ‘verb-Q’ in the embedded clause. They argued that, like the dependency that exists between a displaced wh-filler and its gap position in wh-movement languages, there is a dependency between a wh-word and a related Q-particle in wh-in-situ languages. Due to this dependency, Japanese readers expect a Q-particle as soon as they encounter a wh-word, in the same way that English readers expect a gap as soon as possible after encountering a filler (e.g., Crain and Fodor, 1985; Stowe, 1986; Frazier and Clifton, 1989). Main clause wh-questions were read more slowly than embedded clause wh-questions because they violated this expectation (see also Aoshima et al., 2004).

Building on these prior studies, we used ERPs to gain a clearer view of the temporal dynamics involved in processing the relationship between wh-words in situ and their corresponding Q-particles in mono-clausal and bi-clausal Japanese wh-questions. We hypothesized that there would be a neural signature of the dependency between Japanese wh-elements and their Q-particles, similar to that seen in wh-movement filler-gap dependencies in English or German. In particular, we expected to see slow anterior negative potentials between wh-words and their corresponding Q-particles, similar to those seen between displaced fillers and their

corresponding gaps, and phasic LAN and/or P600 effects at the Q-particle position itself, similar to ERP effects typically seen at gap positions, as schematized in (7).

(7)

- a. Filler-gap dependency for wh-movement languages (e.g., English, German)

- b. Wh-Q dependency for wh-in-situ language? (e.g., Japanese)

‘What did Calvin bring?’

On the other hand, as outlined above, Japanese exhibits a number of different syntactic properties that might elicit different ERP effects in the processing of Japanese wh-questions. For example, since Japanese is a left-branching and strictly head-final language, it is possible that the processing of wh-in-situ questions in Japanese could elicit some other type of ERP effect not evident in the processing of wh-movement questions in English and German, which are right-branching, (mostly) head-initial languages. Moreover, in Japanese the parser has to cope with wh-elements missing a different type of information—namely, interrogative scope—than is the case with English and German wh-fillers, whose grammatical functions and semantic roles are initially unclear, and this could cause an ERP difference as well. In any event, investigating the processing of syntactically distinct languages can help to provide insight into the properties of language-related ERP components, as different types of dependencies in different languages may trigger the same, different, or partially overlapping processes, resulting in the same, different, or partially overlapping ERP effects. This investigation thus holds the promise of revealing language-universal vs. language-specific aspects of the processing of wh-questions, which can lead to better understanding of anterior negativities and posterior positivities in general.

2. Experiment 1

Experiment 1 investigated the processing of mono-clausal Japanese wh-questions as a first step toward the above-stated goals. This investigation was based on Ueno and Kluender (2003), but in this case focused on an entirely different factor of the design, in that it compared the processing of two different types of wh-questions with the processing of their yes/no-question counterparts. Wh-questions included wh-in-situ questions with wh-objects in situ and

scrambled wh-questions with scrambled wh-objects; yes/no-questions included yes/no-in-situ questions with demonstrative objects in situ and scrambled yes/no questions with scrambled demonstrative objects as shown in Table 1 (see Section 5.1.2 for details).

2.1 Results

The mean correct response rate for the probe recognition task across participants was 91% (S.D. 5.2%); thus no participant's data were excluded from the ERP analyses based on poor scores. Comparisons were made both at wh-elements themselves as well as from the words following wh-elements to the sentence-final verb-Q complex, in order to see if there was any neural signature of the wh-Q dependency shown in (7).

2.1.1 Lexical effects at wh-elements—ERPs in response to wh-pronouns ('what-ACC' 何を and 'who-ACC' 誰を)—whether scrambled or in situ—exhibited both larger P200 components and larger and earlier N350 components than the ERPs to demonstrative pronouns ('that-ACC' それ and 'that-person-ACC' あの人は). Figure 1a shows ERPs to wh- vs. demonstrative pronouns collapsed across in situ and scrambled conditions; results of relevant ANOVAs are displayed in Table 2. ANOVAs performed in a latency window of 200 to 300 ms poststimulus onset confirmed that the P200 elicited by wh-words was of greater mean amplitude than the P200 to demonstratives over right frontal regions. Similarly, ANOVAs performed on mean amplitude measures in a latency window of 300 to 375 ms confirmed that the N350 elicited by wh-words was larger than that elicited by demonstratives, especially over the left hemisphere. Additional ANOVAs performed on peak latency (rather than mean amplitude) measures in a window of 250 to 450 ms indicated that the N350 in response to wh-words also peaked earlier than the N350 elicited by demonstratives, especially over posterior regions.

The larger P200 and larger and earlier N350 in response to wh-words initially gave the impression of a wh-Q dependency effect. However, closer examination showed that these differences were instead due to lexical processing differences between wh-words and demonstratives (Figure 1b): significant correlations were found between the P200 and N350 components and lexical factors reported to influence ERPs in these time windows, namely, graphic complexity (Liu et al., 2003) and word length (e.g., Neville et al., 1992; Osterhout et al., 1997; King and Kutas, 1995b, 1998; Hauk and Pulvermüller, 2004). The P200 effect (i.e., larger P200s to wh-words than to demonstratives) was related to the graphic complexity of wh-words, which consisted solely of Chinese characters (*kanji*) in our stimulus set, as these require on average more strokes to write than phonemic (*kana*) representations. Trend analyses (cf. Keppel, 1982) on P200 amplitude (collapsed across electrodes for the four wh- and demonstrative pronouns) in terms of stroke count revealed a significant linear component for total stroke count [$F(1,57) = 14.21, p < .001$]; none of the other higher-order (quadratic, cubic) trend components was significant. Correlations between P200 amplitude and frequency or word length did not reveal any similar significant linear trend component. The N350 mean amplitude effect (i.e., a larger N350 response to wh-words than to demonstratives), on the other hand, was related solely to word length: among stroke count, frequency, and word length, only word length revealed a significant linear trend component by itself [$F(1,57) = 4.76, p = .033$]. Similarly, trend analyses on the N350 peak latency effect (i.e., the N350 to wh-words peaking earlier than that to demonstratives) revealed a significant linear trend component [$F(1, 57) = 6.89, p = .012$] for word length, but not for frequency or stroke count.

In sum, wh-words in comparison to demonstratives elicited a larger P200 component and an earlier, larger N350 component. Mean amplitude of the P200 correlated with graphic complexity as measured by stroke count, whereas both the latency and the amplitude of the N350 correlated solely with word length.

2.1.2 Right-lateralized negativity between wh-elements and Q-particles—To test more directly for brain responses to the dependency between a wh-word and its corresponding Q-particle at sentence end, we also examined ERPs to the words following wh-words up to and including the sentence-final verb-Q complex. The two in situ conditions and the two scrambled conditions (Table 1.a. vs. 1.b. and 1.c. vs. 1.d) were compared separately, as it was assumed that comparing minimal sentence pairs, differing only in the presence of a wh-word, would reveal effects specific to the processing of wh-questions alone.

2.1.2.1 Wh-in-situ questions vs. yes/no-in-situ questions: The wh-in-situ vs. yes/no-in-situ question comparison was a single-word average at the sentence-final verb position ('discovered-Q' in Table 1), as in situ wh- and demonstrative objects immediately preceded the verb (Table 1.a. and 1.b, Figure 2). On visual inspection, wh-in-situ questions appeared to elicit greater broad negativity starting around 300 ms post-stimulus onset of 'discovered-Q', especially over the right side of the head.³ ANOVAs performed within a latency window of 300–600 ms revealed a significant main effect [$F(1, 19) = 9.91, p = .005$] and a marginal condition $\times$ hemisphere interaction [$F(1, 19) = 3.01, p = .099$] within the temporal array. The interaction was due to greater negativity over the right hemisphere in response to wh-in-situ questions. There was no accompanying P600 effect in response to the verb-Q complex in wh-in-situ questions.

2.1.2.2 Scrambled wh-questions vs. scrambled yes/no-questions: In order to compare scrambled wh-questions vs. scrambled yes/no-questions, ERPs to words following scrambled wh- and demonstrative objects up to and including the verb-Q complex ('reckless adventurer-NOM finally discovered-Q' in Table 1.c. and 1.d.)⁴ were examined (Figure 3). Visual inspection of these four-word averages revealed a sustained right-lateralized anterior negativity starting about 300 ms post-onset of 'adventurer-NOM' and continuing through to sentence end in response to the scrambled wh-question condition.

ANOVAs performed in a latency window of 300 ms poststimulus onset of 'adventurer-NOM' to sentence end, including the sentence final verb-Q complex 'discovered-Q' (950 to 2600 ms from the pre-stimulus baseline), showed a significant interaction between condition and hemisphere in the parasagittal array [$F(1, 19) = 6.40, p = .02$], as well as a significant condition $\times$ hemisphere $\times$ anteriority interaction in the temporal array [$F(2, 38) = 3.37, p = .045$]. The interactions were due to the fact that scrambled wh-questions were more negative than scrambled yes/no-questions at right anterior sites. To test the time course of this effect in greater detail, separate ANOVAs were run 300–600 ms poststimulus onset of each of the last three words of the sentence without rebaselining (950–1250 ms for 'adventurer-NOM', 1600–1900 ms for 'finally', and 2250–2550 ms for 'discovered-Q'; see Table 3). The overall pattern remained the same as for the 950–2600 ms analysis, with the exception that the right-lateralized negativity in response to the subject noun 'adventurer-NOM' had a posterior maximum.

³There are three reasons not to treat this difference as a positivity elicited by the processing of an unexpected Q-particle in yes/no-questions. First, the scrambled wh vs. yes/no comparison elicited a difference starting 2–3 words before the verb-Q position (Figure 3). As that difference cannot have been caused by the Q-particle at the end of scrambled yes/no-questions, it is more likely to have been a negativity in response to scrambled wh-questions. It thus seems more parsimonious to assume that the effect in response to the in situ comparison exhibits the same negative polarity. Second, the morphology of the response is unusual for a purported positivity; the divergence in Figure 2 starts before 200 ms poststimulus onset and does not exhibit a reliable positive trough anywhere. Finally, responses to an unexpected Q-particle seem to be realized as N400-like negativity in the present study (Figure 7, fn. 7). On this basis, we maintain that the difference shown in Figure 1 is a negativity to wh-in-situ questions, and proceed with our discussion.

⁴Both sentence types had five words after the scrambled objects: a subject noun phrase consisting of three words, an adverbial, and a verb ('that reckless adventurer-NOM finally discovered-Q' in the sample stimuli). However, since wh-phrases showed lexical P200 and N350 differences relative to demonstratives (cf. Section 2.1.1 and Figure 1), the ERPs to 'what-ACC' and 'that-ACC' were not identical enough to be used as pre-stimulus baselines. Instead, the demonstrative *sono* 'that' within the subject noun phrase was used as the pre-stimulus baseline. This yielded four-word averages of an adjective, a noun, an adverbial, and a verb-Q complex, as in 'reckless adventurer-NOM finally discovered-Q'.

When the averages of the two conditions were re-baselined at the onset of each word (Figure 4), however, there was no evidence of word-by-word increase in negativity. ANOVAs performed in the 300–600 ms window of the subject noun ('adventurer-NOM') showed a marginal main effect of condition [$F(1, 19) = 3.02, p = .098$] in the parasagittal array, a marginal condition $\times$ anteriority interaction in the parasagittal [$F(1, 19) = 2.51, p = .097$] and temporal [$F(1, 19) = 2.95, p = .067$] arrays, and a significant or marginal condition $\times$ hemisphere $\times$ anteriority interaction in the parasagittal [$F(1, 19) = 2.47, p = .053$] and temporal [$F(1, 19) = 4.24, p = .022$] arrays. However, ANOVAs in the 300–600 ms window of the adverbial ('finally') and of the sentence-final verb-Q position ('discovered-Q') revealed no significant or marginal differences. In addition, ANOVAs performed to test for a local P600 (500–800 ms) effect (see (7)) at the sentence-final verb-Q position revealed no significant or marginal differences, either.

To sum up, wh-in-situ questions elicited right-lateralized negativity at the sentence-final verb-Q position, and scrambled wh-questions elicited slow anterior negative potentials, especially over the right side of the head, from the subject noun to the sentence-final verb-Q position. There was no evidence of a word-by-word increase in negativity when the relevant ERPs were rebaselined at the beginning of each word for the scrambled wh-question comparison, nor was there any sign of a local P600 at the sentence-final verb-Q position of wh-in-situ or scrambled wh-questions.

2.2 Discussion

Wh-questions in comparison to yes/no-questions elicited two types of effects in the present experiment. The first involved lexical effects in response to wh-words themselves: compared to demonstratives, wh-words elicited larger P200s (correlated with stroke count) and earlier and larger amplitude N350s (correlated with word length) (Figure 1). The second type of effect was right-lateralized (anterior) negativity between wh-words and Q-particles. Wh-in-situ questions in comparison to yes/no-in-situ questions elicited greater right-lateralized negativity at the sentence-final verb-Q position (Figure 2), while the last three words of scrambled wh-questions (including the verb-Q complex) elicited greater right-lateralized anterior negativity than the same four words in scrambled yes/no-questions (Figures 3 and 4). In what follows, we will discuss each of these effects in turn.

2.2.1 Lexical effects at wh-elements—The interaction between P200 amplitude and stroke count most likely reflects local visual processing of complex Chinese characters, consistent with reports of larger P200s in response to Japanese open-class words, which tend to be written in Chinese characters (*kanji*), than to closed-class words, which tend to be written in phonemic characters (*kana*) (Takashima et al., 2001, 2002; also see Liu et al., 2003 for discussion of the P200 as an index of visual-graphic processing in Chinese). As for the N350, the correlation of word length with earlier peak latency is in part consistent with Osterhout et al. (1997) (who also reported a frequency-based correlation that was not evident in the present study), in that shorter wh-words elicited earlier peaks than longer demonstratives did. The correlation of word length with greater amplitude is in line with Neville et al. (1992), who reported that longer in comparison to shorter open-class words elicited less negative ERPs in a time window following the P200 (Neville et al., 1992: Figure 5), as well as with Hauk and Pulvermüller (2004), who reported that longer open-class words elicited smaller amplitude ERPs in the 150–360 ms window. It is interesting that frequency, which has been reported to exert such a strong influence over lexical ERP effects in English and German (e.g., Van Petten and Kutas, 1990; Osterhout et al., 1997; King and Kutas, 1995b, 1998; Hauk and Pulvermüller, 2004; Dambacher et al., 2006), played no apparent role in the present study.

In terms of the ERP responses alone, the P200 and N350 components can be shown to be independent of each other. Lexically driven N350/N400 effects are at times preceded by less positive P200 responses when the N400 difference begins in the P200 trough (Kluender and Kutas 1993b), but also at times by larger P200s (Neville et al. 1992). In the present study, the larger N350 response to wh-words was preceded by a larger P200 response to the same set of words; therefore it is not the case that the amplitude of the N350 was artificially increased by a smaller preceding P200.

That said, it is true that both the P200 and the N350 are sensitive to lexical factors. To date, studies on the processing of English and German words have focused on word length and/or frequency (e.g., Van Petten and Kutas, 1990; Neville et al., 1992; Osterhout et al., 1997; King and Kutas, 1995b, 1998; Hauk and Pulvermüller, 2004) as crucial variables, while studies on the processing of Chinese characters have focused on graphic complexity, pronunciation, and meaning (e.g., Liu et al., 2003; Hsiao et al., 2007). Our trend analyses revealed separate correlations of graphic complexity (number of strokes) with the P200 and of word length (number of characters) with the N350. These two factors do not seem to be entirely independent of each other: phonemic characters (*kana*) are written with fewer strokes but tend to require more characters per word, whereas ideographic Chinese characters (*kanji*) are written with more strokes but tend to require fewer characters per word. To our knowledge, however, no ERP study before ours has simultaneously manipulated both variables or reported both effects. It is therefore impossible to disentangle the correlation of these two factors with P200 and N350 amplitude (and latency, in the case of N350) using the current experimental design, as our post hoc trend analyses were based on multiple repetitions of only four lexical items, two of each for wh- and non-wh-words. Thus further research will be required to tease apart how stroke counts and number of characters interact with the P200/N350 complex in languages that use both phonemic and ideographic writing systems.

2.2.2 Right-lateralized negativity between wh-elements and Q-particles—The negativity occurring between wh-words and the verb-Q complex did not have an identical distribution in in situ vs. scrambled wh-questions: wh-in-situ questions elicited right-lateralized negativity, while scrambled questions elicited right anterior negativity. However, the two effects had several other aspects in common, namely, question type (elicited by wh-questions), point of occurrence (after processing wh-words), polarity (negative), and lateralization (right-lateralized). In addition, there may be some commonality in the latency of the two effects. For the scrambled comparison, the effect was seen in response to the final few words of the sentence (starting at ‘adventurer-NOM’ for some electrodes) (Figure 3). Although the negativity did not grow word-by-word when the relevant ERPs were rebaselined at the beginning of each word (Figure 4), lack of word-by-word increase in negativity has also been reported for sustained negativity in response to filler-gap dependencies (King and Kutas, 1995; Phillips et al., 2005; Hagiwara et al., 2007) (see Section 3.2.2 for further discussion). For the in situ comparison, there was only one word following the critical word, namely the sentence final verb-Q complex, but the negativity continued across the epoch (Figure 2). Thus both may actually have been sustained rather than phasic effects (we return to this claim in Section 3.2.2). From the above, it seems plausible to consider the two effects to be at least somewhat related to each other. Thus we provisionally treat the two as one type of right-lateralized negativity and proceed with our discussion.

There are several possible ways to interpret this right-lateralized negativity effect. One is as a “sentence-end wrap-up” effect. Previous reports of such effects involve N400-like negativities with posterior distribution at the end of sentences containing syntactic and/or semantic anomalies (Osterhout and Holcomb, 1992; Hagoort et al., 1993; Osterhout and Nicol, 1999; Molinaro et al., 2008), or a widely distributed negativity, also at the end of sentences containing syntactic/or and semantic anomalies (Friederici and Frisch, 2000; Ye et al., 2006). These effects

have been interpreted as showing that the kinds of processing problems that initially evoke N400/P600 effects mid-sentence also make it difficult to integrate the meaning of the entire sentence at its completion. However, unlike studies with anomalies, the stimulus sentences used in the present experiment were all grammatical, and our negativity effects were not preceded by N400/P600 effects earlier in the sentence. In addition, the distribution of the negativity was anterior rather than posterior for at least the scrambled wh- vs. yes/no-question comparison. These points distinguish our effects from the previously reported wrap-up effects in the ERP literature.⁵ Another possibility is that the effects indexed a dispreferred parse. However, to the best of our knowledge, dispreferred parse effects in otherwise grammatical sentences have been linked to the P600 elicited by the critical word (Osterhout and Holcomb, 1992; Osterhout et al., 1994; Friederici et al., 1996) rather than to any sentence-ending negativity, rendering it unlikely that the right-lateralized negativity effects in the present study could be taken as an index of a dispreferred parse.

In view of the unlikelihood of these other possibilities, the effect of right-lateralized (anterior) negativity most likely indexes some type of relationship between Japanese wh-words and corresponding Q-particles. This is consistent with previous experimental claims about Japanese wh-questions, as discussed in Section 1.3. Therefore, the first part of our hypothesis, that there is a neural correlate of the dependency between a Japanese wh-element and its related Q-particle, seems to be well supported.

However, the second part, that the wh-Q dependency is processed in much the same way as a filler-gap dependency, is in need of further consideration. First, reported LAN effects elicited by visual presentation of stimuli involving wh-movement have always been either bilateral or left-lateralized (Kluender and Kutas 1993a, b; King and Kutas, 1995a; Kluender and Münte, 1998; Fiebach et al., 2001, 2002; Felser et al., 2003; Phillips et al., 2005; but see Müller et al., 2007 for right-lateralized anterior negativity reported for an auditory version of King and Kutas, 1995a), but curiously, our effects were right-lateralized. Second, the latency of the effect is unclear. Effects of anterior negativity in response to filler-gap dependencies in English and German have been recognized as long-lasting slow potentials between filler and gap (King and Kutas, 1995a; Kluender and Münte, 1998; Fiebach et al., 2001, 2002; Felser et al., 2003; Phillips et al., 2005), as well as phasic forms at the gap position (Kluender and Kutas, 1993a, b; King and Kutas, 1995a). Recall that we anticipated slow potentials between Japanese wh-elements and their corresponding Q-particles, as well as possible phasic effects at the sentence-final verb-Q position. However, at this point, we cannot tell whether wh-in-situ questions elicited a sustained or a phasic effect, since only the sentence-final word was available for comparison. On the other hand, given that scrambled wh-questions did not elicit any additional effect of phasic negativity at the sentence final verb-Q position when averages were re-baselined at this position, it seems more parsimonious to conclude, at least provisionally, that both effects belong to the family of slow potentials, which may be more related to the continuation rather than completion of a dependency. Lastly, as for the P600, neither the in situ nor scrambled comparison revealed any effect of positivity at the verb-Q position, where the hypothesized wh-Q dependency is presumably closed off.

Given the above, the question is whether the same pattern of results is found not only in monoclausal but also in bi-clausal wh-questions. As shown in (5) above, bi-clausal wh-questions in Japanese provide two possibilities for scope-marking, namely, embedded vs. main clause scope of the wh-question. Experiment 2 thus investigated whether right-lateralized (anterior)

⁵This does not completely eliminate the possibility of a sentence-end wrap-up effect, however, as a number of reading time studies (e.g., Rayner et al., 2000; Hirotsu et al., 2006) have reported clause-end or sentence-end wrap-up effects with no preceding grammatical or semantic violations. Nonetheless, the negativity effects in our study are distinct from previously reported wrap-up effects in the ERP literature, which to our knowledge all involved certain types of violations. Experiment 2 offers better evidence against a sentence-end wrap-up effect account of RAN in response to the final verb-Q complex. See Section 3.2.1.

negativity would be elicited between wh-words and their corresponding Q-particles in biclausal wh-questions as well.

3. Experiment 2

The stimuli used in Experiment 2 were a slightly modified version of Miyamoto and Takahashi’s (2002) reading time study, with the embedded clause scrambled to the beginning of the sentence (note that scrambling is typical although not obligatory for a long embedded clause in Japanese; cf. Yamashita, 2002), as shown in Table 4 (see Section 5.2.2 for details).

As schematized in (8), we predicted that if right (anterior) negativity (hereafter referred to as “RAN”) really indexes a dependency between a *wh*-element and a corresponding *Q*-particle, it should appear between the *wh*-element and the embedded verb-*Q* position in embedded clause *wh*-questions (8a), and between the *wh*-element and the main clause verb-*Q* in main clause *wh*-questions (8b).

(8) Predictions for Experiment 2

a. Embedded clause wh-question

‘Did the accounting manager ask what kind of computer the director bought?’

b. Main clause wh-question

‘What kind of computer did the accounting manager say the director bought?’

Recall that in Experiment 1, there was an evident RAN response to the verb-Q position of wh-in-situ questions when ERPs were baselined at the onset of the final word, but that this was not the case for the verb-Q position of scrambled wh-questions. Although we claimed that it was therefore more parsimonious to treat the RAN effects to both types of questions as slow potentials, we did not offer definitive proof of this. The bi-clausal design of the present experiment was intended to test this claim more closely. Recall further that there was no phasic LAN/RAN effect in scrambled wh-questions—nor was there a P600 effect in either scrambled or in situ wh-questions—in response to the verb-Q position of Experiment 1. The stimulus materials of Experiment 2 allowed us to test once more for possible phasic LAN/RAN and/or P600 effects at equivalent positions, namely, at the embedded verb-Q position for embedded clause wh-questions and at the main verb-Q position for main clause wh-questions.

3.1 Results

The mean correct response rate to the probe recognition task across participants was 87% (S.D. 8.0%). We investigated ERPs at various sentence positions in order to isolate predicted RAN effects, as schematized in Table 5.

3.1.1 Embedded clause object noun phrase—We begin by investigating ERPs in response to the embedded object noun phrase, namely, ‘what.kind of computer-ACC’ or ‘new computer-ACC’ in Tables 4 and 5. For both wh- and yes/no-questions, ERPs were collapsed across embedded and main clause conditions, as words were identical through the embedded object head noun position, ‘computer-ACC’ (Figure 5). In this comparison, visual inspection suggested a sustained anterior negativity in response to all wh-questions. ANOVAs run in the window of 500–1250ms⁶ revealed a significant main effect of condition in the midline [$F(1,19) = 5.18, p = .035$] and parasagittal [$F(1,19) = 6.22, p = .022$] arrays, indicating that all wh-questions were more negative than all yes/no-questions. To examine the scalp distribution of the negativity elicited by both wh-modifiers and the head nouns they modified (‘computer-ACC’), separate measurements were taken in the time windows of 500–700 ms for modifiers and 950–1250 ms (300–600 ms poststimulus onset of ‘computer-ACC’) for head nouns. For the 500–700 ms window (‘what.kind.of’), there was a significant main effect of condition in the parasagittal [$F(1,19) = 5.63, p = .028$] and temporal [$F(1,19) = 10.02, p = .005$] arrays, indicating that wh-questions were overall more negative than yes/no-questions. In addition, there was a marginal condition $\times$ anteriority interaction in the temporal array [$F(2,38) = 2.90, p = .10$], due to greater negativity at frontal sites in response to wh-modifiers. For the 950–1250 ms window (‘computer-ACC’), there was a significant main effect of condition in the midline [$F(1,19) = 5.18, p = .035$] and parasagittal [$F(1,19) = 6.22, p = .022$] arrays because wh-questions were more negative than yes/no-questions.

To summarize, there was an effect of anterior negativity in response to the entire object noun phrase position (wh-modifier + object head noun, ‘what.kind.of computer-ACC’) within the embedded clause of wh-questions. We next turn to the final position of the embedded clause, the clause-final verb, where the embedded and main clause wh-question conditions first diverged.

3.1.2 Embedded verb—The comparison at the embedded verb (‘bought-Q/that’ in Tables 4 and 5) allowed us to test our prediction, based on the results of Experiment 1, that an unresolved wh-Q dependency should elicit right-lateralized (anterior) negativity (RAN). Main clause wh-questions (Table 4.a.) had no Q-particle attached to the verb of the embedded clause to close off the scope of the wh-phrase and limit it to the embedded clause, while embedded wh-questions (Table 4.b.) did. There was no wh-Q dependency in the yes/no-question control (Table 4.c.) at the embedded verb position. In this comparison, then, there should be an unresolved dependency at the embedded verb position of main clause wh-questions (Table 4.a.), whereas the dependency should be resolved at this point in embedded wh-questions (Table 4.b.), and there should be no such dependency present at all in main clause yes/no-questions (Table 4.c.). Thus the unresolved dependency of main clause wh-questions was predicted to elicit RAN relative to the resolved (embedded clause wh-questions) or nonexistent dependencies (main clause yes/no-questions) of the other two conditions. In addition, the

⁶There was also an N400 effect in response to non-wh-modifiers. This was probably because 50% of wh-modifiers were ‘what.kind.of’ (with adjectives such as ‘new’ as non-wh counterparts), while the other 50% were numeral classifiers such as ‘what-cup-of (how many cups of)’ (with ‘2-cup-of’ as non-wh counterparts, see Section 5.2.2). The N400 amplitude in response to ‘what.kind.of’ seemed to have been *reduced*, yielding an N400 effect to non-wh modifiers. This was attributable to the high rate of repetition (25% [50/200] of experimental sentences; Van Petten et al., 1991), high frequency of occurrence in the language (a mean frequency of 27,715 vs. 6,497 for the non-wh adjectives in Experiment 2; Amano and Kondo, 2000; Van Petten and Kutas, 1991), closed-class status (Münste et al., 2001), and low semantic content (Kluender and Kutas, 1993b; McKinnon and Osterhout, 1996) of ‘what.kind.of’. In order to avoid this N400 effect, the time window for measurement was moved to 500–1250ms.

comparison between embedded clause wh- (Table 4.b.) and yes/no-questions (Table 4.d.) at this sentence position allowed us to test for local LAN/RAN and P600 effects when the scope of embedded clause wh-questions was disambiguated.

Figure 6 shows ERPs to main clause wh-questions, embedded clause wh-questions, and main clause yes/no-questions at the embedded verb position. On visual inspection, when compared to both main clause yes/no-questions and embedded clause wh-questions, main clause wh-questions appeared to elicit a RAN effect that continued throughout the present epoch and into the next, as predicted. For the main clause wh- vs. yes/no-question comparison, ANOVAs performed in the 300–600 ms window showed no significant or marginal differences, while ANOVAs in the 300–900 ms window (i.e., through to the end of the epoch) revealed a significant condition $\times$ anteriority interaction in the midline array [$F(2, 38) = 3.74, p = .035$], due to greater frontal negativity in response to main clause wh-questions than to main clause yes/no-questions. There was also a marginal condition $\times$ hemisphere interaction in the parasagittal [$F(1, 19) = 3.36, p = .083$] and temporal [$F(1, 19) = 3.67, p = .071$] arrays, due to the fact that main clause wh-questions were more negative than main clause yes/no-questions over the right hemisphere. Since this RAN effect appeared more pronounced toward the end of the epoch, separate ANOVAs were run in the 600–900 ms window for this comparison. This resulted in a significant or marginal condition $\times$ anteriority interaction in all three arrays [*midline*: $F(2, 38) = 6.94, p = .003$; *parasagittal*: $F(4, 76) = 4.43, p = .02$; *temporal*: $F(2, 38) = 3.10, p = .091$], as well as a significant or marginal condition $\times$ hemisphere interaction in the parasagittal [$F(1, 19) = 4.33, p = .051$] and temporal [$F(1, 19) = 4.86, p = .04$] arrays, indicating that the effect became stronger in the later time window, consistent with the visual impression.

For the main clause wh- vs. embedded-clause wh-question comparison, ANOVAs performed in the 300–600 ms window again showed no significant or marginal differences, but ANOVAs in the 300–900 ms window revealed a marginal condition $\times$ anteriority $\times$ hemisphere interaction in the temporal array [$F(2,38) = 2.88, p = .071$]. This interaction was because main clause wh-questions in comparison to embedded clause wh-questions were more negative at right anterior/central sites. Although the effect was statistically weak, separate measurements run in the 600–900 ms window showed improved results, just as was found in the main clause wh- vs. yes/no-question comparison above. There was a marginal condition $\times$ anteriority interaction [$F(2,38) = 3.27, p = .084$] and a significant condition $\times$ anteriority $\times$ hemisphere interaction [$F(2, 38) = 3.29, p = .05$] in the temporal array, indicating that the effect again became stronger in the later time window.

ERPs to embedded clause wh- vs. yes/no-questions were also compared at the embedded verb position. There were no phasic LAN/RAN or P600 effects apparent on visual inspection, or evidenced by ANOVAs performed in the 300–600 ms and 300–900 ms windows to detect additional phasic LAN/RAN effects, or by ANOVAs in the 500–800 ms window to detect P600 effects.

To summarize, main clause wh-questions elicited a RAN effect relative to both main clause yes/no-questions and embedded clause wh-questions beginning at the embedded verb position and continuing into the next word. As predicted, this effect seems related to the presence of an unresolved wh-Q dependency in main clause wh-questions. Comparisons of embedded clause wh-questions with embedded clause yes/no-questions revealed no additional phasic LAN/RAN or P600 effects in response to embedded wh-questions, consistent with the results for scrambled mono-clausal wh-questions in Experiment 1.

3.1.3 From embedded to main clause verbs—In the previous section, we discussed RAN effects at the embedded verb position—where wh-question conditions first diverged from each other—that were related to the presence of an unresolved wh-Q dependency in main clause

wh-questions. These effects began at the embedded verb position in accordance with our predictions, but became more pronounced toward the end of the epoch and appeared to continue into the next word. If it is the case that RAN is elicited by the presence of an unresolved wh-Q dependency, then the RAN effect seen in response to main clause wh-questions at the embedded verb position should in fact continue throughout the ensuing main clause region until the scope-marking Q-particle is encountered at sentence end. Thus far in Experiment 2, the RAN-like effects that we have seen in response to entire wh-object noun phrases (Figure 5), and at the embedded verb of main clause wh-questions (Figure 6) suggest that this should be the case. Moreover, the comparison shown in Figure 3 of Experiment 1, in which a wh-phrase was separated from its scope-marking Q-particle by means of scrambling, suggests that this continued RAN effect in the main clause should manifest as a slow right-lateralized anterior negative potential.

With this prediction in mind, we now turn to an examination of the latter half of the stimulus sentences, namely from the embedded verb to the main clause verb ('bought-that/Q accounting-of manager-NOM said/asked-Q' in Tables 4 and 5). In this sentence region, a RAN effect should be evident only in main clause wh-questions (Table 4.a.), as this is the only condition in which an unresolved wh-Q dependency still exists in the main clause. Embedded clause wh-questions (Table 4.b.) have presumably already closed off this dependency at the embedded verb-Q position, and there should be no such syntactic need for a Q-particle in the main clause yes/no-question control condition (Table 4.c.). Thus comparisons were made between main clause wh-questions vs. main clause yes/no-questions, and between main clause wh-questions vs. embedded clause wh-questions in a latency window that covered the last four words of the sentence ('bought-Q/that accounting-of manager-NOM asked/said-Q'), i.e., from the embedded verb through the sentence-final main clause verb.

Figure 7 shows the ERPs in response to these comparisons. A sustained anterior negativity in response to main clause wh-questions was observed in comparison with main clause yes/no-questions. ANOVAs run in the 300–1950 ms window (from 300 ms poststimulus onset of 'bought-that' to the end of 'manager-NOM')⁷ revealed a significant or marginal condition x

⁷ANOVAs run in a latency window of 300–2600 ms revealed a significant condition x hemisphere interaction [$F(1,19) = 4.59, p = .043$], as well as a marginal condition x anteriority interaction [$F(2,38) = 3.76, p = .063$] in the temporal array. These interactions were due to the fact that main clause wh-questions in comparison to main clause yes/no-questions elicited greater negativity more prominently over right frontal electrodes. However, this comparison was contaminated by the presence of an N400-like effect at sentence end in response to main clause yes/no-questions. This N400-like effect is indicated by the dotted arrows in Figure 6, as well as by the greater positivity seen in the "main wh – main yes/no" isovoltage map of Figure 7 in the 2250–2550 ms latency window (i.e., 300–600 ms post-stimulus onset of 'said-Q'). When the final word of the sentence eliciting this N400-like effect in response to yes/no-questions was excluded from the average, the RAN effect to main clause wh-questions emerged more clearly.

As for the statistics of the N400 effect, when the main clause wh- and yes/no-question conditions were re-baselined and compared at this sentence-final position, ANOVAs run in the latency window of 300–500 ms showed a marginal condition x hemisphere interaction in parasagittal [$F(1,19) = 4.01, p = .06$] and temporal [$F(1,19) = 3.84, p = .065$] arrays, as well as a significant condition x hemisphere x anteriority interaction in the parasagittal array [$F(4, 76) = 3.77, p = .008$]. These interactions were due to the fact that main clause yes/no-questions were more negative than main clause wh-questions over left central regions of scalp. However, although the relative effect was larger on the left, the absolute values of the negative peak in response to main clause yes/no-questions were larger on the right, in accordance with the standard distribution of the N400. In other words, since main clause wh-questions were more negative over the right side of the head (due to the RAN effect), the relative difference with main clause yes/no-questions turned out to be larger over the left side of the head.

This N400-like effect is unlikely to be due to the word length or frequency of the relevant verbs, as all the main verbs were identical across conditions. Instead, this is likely to be an effect of seeing an unexpected Q-particle in a sentence that otherwise could have ended as a declarative. That is, if there is no Q-particle at the end of a main clause yes/no-question, such as *Senmu-ga atarashii pasokon-o katta-to keiri-no kakaricho-ga iimashita-ka* 'Did the accounting manager say that the director bought a new computer?', the sentence would end as a declarative, as in *Senmu-ga atarashii pasokon-o katta-to keiri-no kakaricho-ga iimashita* 'The accounting manager said that the director bought a new computer.' Although there were many more questions than declaratives in the stimulus sentences (275 (79%) out of 350), participants still might have expected declaratives by default, by virtue of their life experience, and/or by the fact that only 37% (137 out of 350) of the stimulus sentences were yes/no-questions.

⁸In addition to the N400-like effect discussed in fn 7, the four-word average also had a high rate of blinks at the sentence-final word ('said-Q') that skewed the data for the entire four-word epoch even after blink correction. On account of this, the three-word average file was used for computations of this time window in order to maximize signal-to-noise ratio.

anteriority interaction in the midline [$F(2,38) = 3.87, p = .032$] and temporal [$F(2,38) = 2.51, p = .098$] arrays, in addition to a significant condition $\times$ hemisphere interaction in the temporal array [$F(1,19) = 4.7, p = .043$]. These interactions were due to the fact that main clause wh-questions in comparison to main clause yes/no-questions elicited greater negativity more prominently over frontal electrodes or over the right hemisphere. In order to investigate the time course of this effect more precisely, separate ANOVAs were run 300–600 ms poststimulus onset of each word without re-baselining (300–600 ms for ‘bought-that’, 950–1250 ms for ‘accounting-of’, 1600–1900 ms for ‘manager-NOM’, and 2250–2550 ms for ‘said-Q’, see Table 6). Although the results became weaker, the significant/marginal interactions overall pointed to right-lateralized anterior negativity in response to main clause wh-questions.

When the averages of the two conditions were re-baselined at the onset of each word (Figure 8), however, ANOVAs performed in the 300–600 ms window of each word revealed no evident increase in negativity, just as was seen in Experiment 1 (Figure 4). There was a marginal condition $\times$ hemisphere $\times$ anteriority interaction in the parasagittal array [$F(4, 76) = 2.12, p = .089$] at the sentence final verb (‘said-Q’), but this was because main clause yes/no-questions were more negative than main clause wh-questions, which appeared to be due to an N400-like effect in response to main clause yes/no-questions (see fn. 7). Because of this N400-like effect, additional ANOVAs were run in the 300–900 ms window of the sentence-final verb to test the existence of phasic negativity in response to main clause wh-questions, but they revealed no significant or marginal differences. ANOVAs performed in the 500–800 ms window of the sentence-final verb (to test for a P600 effect) did not reveal any significant or marginal differences, either.

When we switch our comparison to main clause wh-questions (Table 4.a.) vs. embedded clause wh-questions (Table 4.b.), a similar sustained anterior negative response to main clause wh-questions was observed (Figure 7). This impression was supported by ANOVAs run in the 300–2600 ms window (a longer latency window was possible in this case because there was no confounding N400 effect at the sentence-final verb; see preceding paragraph). There was a significant or marginal main effect of condition in the parasagittal [$F(1,19) = 3.2, p = .09$] and temporal [$F(1,19) = 4.71, p = .043$] arrays, as well as a marginal condition $\times$ anteriority interaction in the midline array [$F(2,38) = 2.98, p = .095$]. As in the previous main clause wh- vs. yes/no-question comparison, these results were due to the fact that main clause wh-questions elicited greater negativity than embedded clause wh-questions through the end of the sentence, especially at frontal electrodes. In order to investigate the time course of this effect more precisely, separate ANOVAs were again run 300–600 ms poststimulus onset of each word without baselining (Table 6). The results indicated the existence of anterior negativity in response to main clause wh-questions across the epoch, with a focus on the final word. It should also be noted that the marginal condition $\times$ hemisphere $\times$ anteriority interaction in the temporal array in the 950–1250 ms window was due to the fact that main clause wh-questions were more negative especially over right frontal regions.

Similar to the main clause wh vs. yes/no-question comparison discussed above, however, re-baselining ERPs at the onset of each word revealed no evident successive increase in negativity (Figure 8). ANOVAs performed in the 300–600 ms window of each word revealed a significant condition $\times$ hemisphere $\times$ anteriority interaction in the temporal array [$F(2, 38) = 5.38, p = .001$] at the subject noun modifier (‘accounting-of’), which was due to greater negativity mainly at right frontal regions in response to matrix-wh questions, but there were no other significant or marginal differences at any sentence position. Additional ANOVAs performed in the latency windows of 300–500 ms (N400), 300–900 ms (LAN/RAN), or 500–800 ms (P600) at the sentence-final verb revealed no significant or marginal differences, either.

In summary, main clause wh-questions elicited right-lateralized anterior negativity when compared to main clause yes/no-questions, and anterior negativity with some evidence of right-lateralization when compared to embedded clause wh-questions, in the four-word window extending from the embedded verb to the sentence-final main clause verb-Q position. However, rebaselining the corresponding ERPs at the beginning of each word revealed no evidence of word-by-word increase in the negativity effects, nor any additional phasic LAN/RAN or P600 effects at the sentence-final main verb-Q position.

3.2 Discussion

To summarize, we have seen effects of anterior negativity at various positions between wh-words and corresponding Q-particles of wh-questions in multiple comparisons (Figures 5–8). The effects found in the present bi-clausal experiment (and the mono-clausal Experiment 1) were right-lateralized in the majority of cases, although they were sometimes bilaterally distributed. These sustained anterior negativity effects, however, showed no evidence of word-by-word increase when the corresponding ERPs were re-baselined at the beginning of each word, just as was the case in Experiment 1. In addition, there were no local LAN/RAN or P600 effects at either the embedded verb-Q or main verb-Q positions of the corresponding wh-questions.

Recall that the remaining questions from Experiment 1 were (a) whether RAN effects would be elicited between wh-words and their corresponding Q-particles for bi-clausal wh-questions with embedded and main clause scope, (b) whether such effects would be phasic or sustained, and (c) whether there would be really no local LAN/RAN or P600 effects at the scope-marking verb-Q positions. We address each question below in turn.

3.2.1 RAN in response to wh-Q dependencies—In the present experiment, both main clause and embedded clause wh-questions elicited RAN effects from the wh-modifier position in the embedded clause to the relevant Q-particle position, whether at the end of the embedded clause for embedded clause wh-questions or of the main clause for main clause wh-questions. As shown in Figure 5, all wh-questions in comparison with all yes/no-questions elicited an effect of anterior negativity at the embedded object noun phrase region. As shown in Figures 6–8, main clause wh-questions in comparison to main clause yes/no-questions and embedded clause wh-questions elicited (right-lateralized) effects of anterior negativity. These effects began at the wh-phrase (the onset of the purported wh-Q dependency) and persisted up to the corresponding Q-particle (the end of the wh-Q dependency), providing supporting evidence that they are correlated in some way with the presence of a wh-Q dependency. In addition, Experiment 2 provided stronger evidence against a “sentence-end wrap-up effect” account of the RAN effects seen in Experiment 1. There were (R)AN effects not only at sentence end, but also in response to words in the embedded clause starting at the wh-modifier position, five words before the final word of the sentence. Therefore, we can conclude that (R)AN effects are elicited between wh-words and their corresponding Q-particles in not only mono-clausal but also bi-clausal wh-questions, and that they seem to correlate with the dependency between a wh-word and its corresponding Q-particle.

3.2.2 Latency of RAN effects—Experiment 2 clearly showed that the RAN effects elicited in the experiment were slow potentials, as the effect started at the wh-modifier position and continued up to the relevant verb-Q position. This suggests that the right-lateralized negativity at the verb-Q position of mono-clausal wh-in-situ questions in Experiment 1 was elicited because the wh-phrase immediately preceded the verb-Q position, rather than as a response to the verb-Q position itself.

However, one difference between the two experiments was that the RAN effects elicited in Experiment 2 were apparent across the entire sentence, while those elicited by scrambled wh-questions in Experiment 1 were seen only towards sentence end. This difference cannot be attributed to overall differences in the distance between a wh-element and its corresponding Q-particle across the two experiments, as the mean number of words was 2.4 between wh-phrases and Q-particles collapsed across embedded and main clause wh-questions in the bi-clausal Experiment 2, while the mean distance in the mono-clausal Experiment 1 was 3 words. Instead, it may have something to do with the proportion of bi-clausal wh-questions in the stimulus materials of the two experiments. In Experiment 2, 73% (100 out of 137) of wh-questions were bi-clausal; the only mono-clausal wh-questions were those of a single filler condition. On the other hand, in Experiment 1, only 20% of wh-questions (one filler condition: 25 out of 125 wh-questions) were bi-clausal. Thus when participants saw wh-elements in Experiment 2, they may have immediately expected a biclausal structure. Participants may have anticipated that some of the wh-Q dependencies (for main clause wh-questions) were going to cross a clause boundary, and the anticipation of this operation may have in turn triggered an earlier onset of the RAN effect (see Frazier and Clifton, 1989 for increased processing costs for crossing a clause boundary, and Kluender and Münte, 1998 for a similar result in German wh-questions).

Another point to note is that our RAN effects do not increase at successive word positions. Recall that when ERPs in response to main clause wh-questions were rebaselined at the beginning of each word, we did not find any evident word-by-word increase in negativity (Figure 8). This was true for ERPs in response to scrambled wh-questions (Table 1.c.) in Experiment 1 (Figure 4) as well. This is moreover consistent with the findings of several other ERP studies that have reported sustained negativity between fillers and their gaps with no accompanying word-by-word increases (King and Kutas, 1995: Figure 3; Phillips et al., 2005: Figures 4 and 5; Hagiwara et al., 2007). If one assumes that increasing working memory load should be reflected in a word-by-word increase in negativity, the lack of such an increase in these studies seems at odds with a number of behavioral results that have clearly demonstrated the parser's preference for shorter wh-dependencies (e.g., Stowe, 1986; Frazier and Clifton, 1989; de Vincenzi, 1991; also Experiment 1 of Phillips et al., 2005). It also seems to conflict with reports of larger sustained negativity in response to longer vs. shorter dependency conditions within the same experiment when across-sentence averages are used (Fiebach et al., 2001, 2002; Hagiwara et al., 2007).

How can one solve this paradox? It could be an artifact of rebaselining, in that the brain activity in response to the previous word when used as a prestimulus baseline may artificially mask the effect in response to the current word. However, one study reported no word-by-word increase in negativity even when a post-stimulus baseline was used in order to avoid this potential confound (Phillips et al., 2005). In fact, even with the same post-stimulus baseline used in Phillips et al., the present data still do not reveal a word-by-word increase in negativity, suggesting that the nonexistence of a word-by-word increase is unlikely to be due solely to baseline artifacts.

A more promising account is to postulate that there is a steady level of continuous brain activation indexing a sustained process throughout the course of a wh-dependency. The connection between working memory load and sustained processes has been supported by a wide range of data from single cell recording studies with monkeys (e.g., Goldman-Rakic, 1987; Fuster, 1997) to ERP studies with humans involving retention (Ruchkin et al., 1990, 1992) and retrieval and evaluation (Rösler et al., 1993, 1995) of information. In terms of ERPs, Rösler et al. (1993, 1995) in particular showed that the negativity for a given memory load is fairly constant over time, thus suggesting that there can be constant activity while the load is maintained. In other words, we can conclude that as long as ERPs have not returned to the

default state, which is clearly the case in multi-word time windows (King and Kutas, 1995: Figure 1; Phillips et al., 2005: Figures 4 and 5; Hagiwara et al., 2007: Figure 2; present study: Figures 3 and 7), the sustained process has not ended, although it may not be increasing.

It also seems feasible to think that a sustained process over a longer time window involves a higher cumulative processing cost. For instance, if we suppose that the cost of carrying a wh-filler for a sustained process is continuously one at each sentence position (compared to a structurally and otherwise lexically matched non-wh control sentence), carrying a filler for a longer distance (9b) than a shorter distance (9a) would incur a higher cumulative cost.

(9)

a. Shorter dependency

cost of carrying a filler	cumulative cost
1 1 1 1	1+1+1+1=4
What would he want __ for his birthday?	

b. Longer dependency

1 1 1 1 1 1 1	1+1+1+1+1+1+1=7
What do you think he would want __ for his birthday?	

In this way, a reasonable explanation is available for why sustained anterior negativity effects do not grow at successive word positions, while the parser seems to be at the same time sensitive to the length of a wh-dependency, in both the present study (wh-Q dependencies) and other studies (filler-gap dependencies).

3.2.3 Wh-Q resolution—no local LAN/RAN or P600 effect?—Just as in Experiment 1, Experiment 2 revealed no local LAN/RAN or P600 effect at the verb-Q positions where wh-scope was disambiguated, either at the embedded verb-Q position for embedded clause wh-questions or at the main verb-Q position for main clause wh-questions. We conclude that this is because there is no physical displacement of wh-fillers in Japanese wh-questions (see Section 4.2).

While the results of Experiment 2 were in general consistent with those of Experiment 1 in terms of no local LAN/RAN or P600 effects, there is one respect in which they were not entirely equivalent. Mono-clausal wh-in-situ questions vs. yes/no-in-situ questions in Experiment 1 elicited right-lateralized negativity at the sentence-final verb-Q position (Figure 2). The same was not true of any of the other comparisons of wh-questions to their yes/no-question counterpart controls at final verb-Q positions in either experiment: scrambled wh-questions vs. scrambled yes/no-questions (Table 1.c. vs. 1.d.) in Experiment 1, as well as embedded wh-questions vs. embedded yes/no-questions, or main clause wh-questions vs. main clause yes/no-questions in Experiment 2 as discussed above.

The question then is why the final-word averages of monoclausal wh-in-situ questions (Table 1.a.) in Experiment 1 exhibited this minority pattern of pronounced RAN to a sentence-final verb-Q complex. First, we suspect that the respective positions of wh-elements and Q-particles in our stimulus materials contributed to this difference. In scrambled wh-questions (Table 1.c.) and main clause wh-questions (Table 4.b.), wh-elements triggering the dependency occurred five words away from the sentence-final point of comparison. Although the embedded verb in embedded clause wh-questions (Table 4.a.) was technically preceded by the entire wh-phrase, the actual wh-modifier itself occurred two words back. In mono-clausal wh-in-situ questions (Table 1.a.), on the other hand, the sentence-final verb was immediately preceded by a bare

wh-phrase. Second, the RAN effects to both scrambled and in situ mono-clausal wh-questions in Experiment 1 were significantly delayed (Figures 2 and 3) relative to the RAN effects in Experiment 2, which were immediate (Figure 5). In section 3.2.2 we suggested that this delay may have been a result of strategic processing on the part of participants, due to the greater proportion of bi-clausal stimuli in Experiment 2 (73%) vs. Experiment 1 (20%), which caused participants to anticipate wh-Q dependencies across a clause boundary (main clause wh-questions) in Experiment 2. In any event, because the wh-word in the wh-in-situ condition immediately preceded the sentence-final verb-Q complex (Table 1.a.), the delay in this condition manifested exceptionally as a RAN effect on the subsequent word, which in this case just also happened to be the final word of the sentence.

4. General Discussion

Taken together, the two experiments showed that both mono-clausal and bi-clausal Japanese wh-questions elicited right-lateralized anterior negativity (RAN) between wh-words and corresponding Q-particles relative to their yes/no-question counterparts. This pattern of results is similar to ERP effects seen between wh-fillers and gaps in English and German, and suggests a reliable neural processing correlate of the dependency between Japanese wh-elements and Q-particles, but with a right- rather than left-lateralized distribution. We address some related issues below.

4.1 LAN vs. RAN

We have seen that both filler-gap dependencies and wh-Q dependencies elicit anterior negativity, often in the form of slow potentials. As discussed in Section 1.1, filler-gap dependencies have been reported to elicit sustained left anterior negativities between filler and gap, which is said to index increased working memory load due to the parser's purported need to store the filler until its gap position (King and Kutas, 1995a; Kluender and Münte, 1998; Fiebach et al., 2001, 2002; Felser, et al., 2003; Phillips, et al. 2005). The linkage between sustained LAN and working memory load is independently supported by reports of load-sensitive sustained (left) anterior negativities caused by retention (Ruchkin et al., 1990, 1992) and retrieval and evaluation (Rösler et al., 1993, 1995) of verbal information, as well as reports of correlations between the amplitudes of sustained LAN and reading memory span (Daneman and Carpenter, 1980) for sentential stimuli without filler-gap dependencies (Münte et al., 1997; Vos et al., 2001). In addition, some studies of the processing of filler-gap dependencies have reported correlations between reading span and reading times (King and Just, 1991), between reading span and the distribution and amplitudes of sustained LAN (Fiebach et al., 2001, 2002), and between comprehension scores (which are assumed to be linked to reading span) and the amplitudes of sustained anterior negativity (King and Kutas, 1995), although the exact nature of these correlations is not entirely clear. Therefore, we maintain that the processing of filler-gap dependencies involves working memory load as indexed by sustained LAN effects.

Furthermore, we consider our RAN effects analogous to these LAN effects for several reasons. First, although our effects were often right-lateralized, they were always found bilaterally and never solely over the right side of the head. In addition, it should be kept in mind that ours is neither the first nor the only study to report bilateral and/or right-lateralized slow negative potentials in response to syntactic dependencies. Although anterior negativities have been found to be left-lateralized to varying degrees in several studies (hence the name LAN), there have been several observations of bilateral distributions (see Vos et al., 2001 for a review of variation in the lateralization of anterior negativities across studies). In fact, almost all of the reports of sustained LAN effects discussed in relation to filler-gap dependencies (King and Kutas, 1995a; Münte and Kluender, 1998; Fiebach et al., 2001, 2002; Phillips et al., 2005)

show a bilateral negativity which may be larger on the left, and right-predominant anterior negativity with auditory presentation (e.g., Müller et al., 1997) has also been reported. Therefore, it seems feasible to interpret the sustained anterior negativities found in the present study as an index of working memory load due to the dependency between a wh-element and its corresponding Q-particle in Japanese, as the parser would have to keep track of an ongoing unresolved wh-dependency until the corresponding Q-particle. However, the question remains: why RAN instead of LAN?

One may think that since the Japanese writing system uses many Chinese characters (*kanji*), which are ideographic and tend to be graphically complex (see Section 2.2.1), the processing of a long-distance dependency written with *kanji* in Japanese might cause right-lateralization. In fact, there have been a number of reports of right-hemispheric advantages in *kanji* processing (e.g., Nakagawa, 1994; Yamaguchi et al., 2002; Nakamura et al., 2005), and it may even be possible to link the hypothesized *kanji* processing effect to reports of sustained R(A)N for retaining visual images (Barrett and Rugg, 1989; Mecklinger, 1998). However, comparisons between scrambled and canonical word order sentences in Japanese (which presumably involve the same processing operations as filler-gap dependencies) have been reported to elicit (L)AN instead of RAN (Ueno and Kluender, 2003; Hagiwara et al., 2007; see also Kwon 2008 for a similar result with Korean scrambling). Most importantly, the stimulus sentences used in Ueno and Kluender (2003) were identical to those used in Experiment 1 of the present study (but different combinations of the experimental conditions were compared), with the same set of *kanji*. Thus it is not very likely that all Japanese dependencies elicit RAN due to the use of *kanji*; rather, the lateralization difference seems specifically linked to certain properties of Japanese wh-questions.

A further possibility is a difference in the type of missing information that the parser has to deal with for wh-movement vs. wh-in-situ questions. As discussed in Section 1.1, LAN effects in wh-movement languages have been tied to the necessity of holding a wh-filler in working memory without an assigned grammatical function and/or semantic role until a corresponding gap position can be located that will provide this critical information. By way of contrast, a Japanese wh-element is ambiguous not with regard to grammatical function or semantic role, since it is case-marked and located in an appropriate syntactic position, but with regard to its interrogative scope. Whereas a wh-filler needs to be linked with a corresponding gap, a wh-phrase in situ needs to be linked to a corresponding Q-particle. Thus for both wh-movement and wh-in-situ languages, the parser must maintain an unresolved dependency for some amount of time, but the type of information necessary to resolve the dependency differs. It is syntactic for a displaced wh-phrase (namely, the appropriate grammatical function and semantic role—although the latter is of course also partly semantic) but logico-semantic for a wh-phrase in situ (interrogative scope). The computation of syntactic vs. logico-semantic properties could involve different neural generators that contribute to the left- vs. right-lateralization differences. However, although scope ambiguity seems to be the driving force for wh-Q dependencies, to the best of our knowledge there is no independent evidence to link logico-semantic processes to right-lateralization, as logical or mathematical processes have been linked to phasic or sustained posterior/parietal negativities in terms of ERPs (Ruchkin et al., 1991; Martín-Loeches et al., 2006), left hemisphere activation in terms of MEG (Yoshida et al., 1998), and inferior parietal and lateral frontal activation in terms of PET and fMRI (cf. Feh et al., 2007). Thus the “missing information type” account of the lateralization differences remains tenuous.

A somewhat related alternative explanation is an implicit prosody effect. Although stimulus sentences were presented visually in the present study, there have been reports of the influence of implicit prosody in ambiguity resolution with self-paced reading times (Bader, 1998; Hirose, 1999; Fodor, 2002). Participants may have been saying sentences to themselves silently or

imagining corresponding prosody while reading sentences (more likely with an appropriate presentation rate for doing so; see Section 5.1.3), and this could cause a tacit prosody effect. In fact, there is a well-known correlation between prosody and the scope of wh-questions in different dialects of Japanese. In the Tokyo dialect this manifests as a high pitch on the wh-element itself, followed by a sustained de-accented low pitch on all words within its scope (i.e., the lexically specified pitch accents of all the words within the interrogative scope are significantly attenuated), usually up to the corresponding Q-particle position (e.g., Ishihara, 2002; Deguchi and Kitagawa, 2002; Hirotsu, 2003). A pitch analysis of a spoken version of our sample stimuli (Ueno, 2003) suggests that our wh-questions follow this pattern as well, while there was no such pitch reduction in the yes/no-question counterparts. A number of studies report that the right hemisphere predominates in interpreting prosodic cues (e.g., Zatorre et al., 1992; Meyer et al., 2002), and there have been several reports of phasic ERAN/RAN effects for prosody mismatches both in terms of music (e.g., Patel et al., 1998; Koelsch et al., 2000; Koelsch and Jentschke, 2008) and of sentential contour (e.g., Eckstein and Friederici, 2005, 2006). Thus we could link the right lateralization of the negativity in our data to an implicit prosody effect, in that participants may have been imagining the pitch contour specific to Japanese wh-questions within the domain of their scope. This may have caused more involvement of the right hemisphere and the subsequent right-lateralization of anterior negativity in response to our stimulus materials.⁹

However, one could just as well claim that participants were imagining a lower pitch in the scope of a wh-element because they had not yet located the required Q-particle in the sentence string. In this sense, an effect of implicit prosody could be subsumed under the effect of the unresolved wh-Q dependency as part of the same scope disambiguation process.

4.2 Lack of phasic LAN/RAN and P600 effects

Recall that both experiments revealed no local LAN/RAN or P600 effect at the verb-Q positions at which wh-scope was disambiguated. This suggests that physical displacement of a filler is necessary to elicit such effects. Recall that phasic LAN effects have been tied to the retrieval of a filler from working memory (Kluender and Kutas, 1993a,b; King and Kutas, 1995) or the parser's backward search for an appropriate filler (Kwon, 2008), and that P600 effects have been tied to the syntactic integration of wh-fillers (Kaan et al., 2000). Thus the simplest explanation for the lack of such effects in Japanese wh-questions is that Japanese wh-words remain in situ. In other words, unless a sentence constituent is physically displaced, there is nothing to retrieve or syntactically integrate later in the sentence.

It has to be kept in mind that it is not the case that the Japanese language never elicits phasic LAN or P600 effects. When a wh/non-wh object is displaced via scrambling in Japanese, similar ERP effects to those for filler-gap dependencies in English/German have been reported. These effects include sustained LAN between the object filler and its gap, as well as phasic P600 and LAN effects around the gap position itself (Ueno and Kluender, 2003; Hagiwara et al., 2007); this suggests that the same parsing operations of holding an object filler in working memory across the course of a sentence and retrieving and integrating it at the gap position apply to Japanese. Thus the lack of phasic LAN and P600 effects in the present study is most likely due to the fact that Japanese does not displace wh-phrases in wh-questions.

⁹This must remain a speculation on our part, because it is impossible to infer source generators from the scalp distribution of ERP effects, and we were unable to perform current source density analyses given the limited number of electrodes used in this experiment. We feel that this speculation is nonetheless warranted, however, given prior research indicating left hemisphere sources for LAN and verbal working memory effects in general (Ruchkin et al., 1990, 1992). We merely assume that the additional influence of prosodic processing in these experiments may have skewed the scalp distribution of the anterior negativity to the right.

4.3 Processing of Japanese wh-questions

Given the present data and previous related studies, it seems fairly clear that Japanese wh-questions incur processing costs that are related to an unresolved dependency between a wh-element and its corresponding Q-particle, both in terms of longer reading times (Miyamoto and Takahashi, 2002; Aoshima et al., 2004) and sustained RAN (present study). We have argued above that sustained RAN seems to index working memory load, and if it does, this would be related to the fact that there is an ongoing unresolved wh-dependency for which scope has not been determined, and that the right-lateralization could be due to an implicit prosody effect. While there was no phasic LAN or P600 effect at the end of a given wh-Q dependency, this seems to be due to the in situ property of Japanese; since Japanese does not displace wh-phrases, there is nothing to retrieve or syntactically integrate at the Q-particle position. Based on the above observations, we propose that Japanese wh-questions are processed in the following way.

Upon recognizing a wh-element (with missing scope information), the parser encodes the existence of an unresolved wh-Q dependency (not so much as a filler, since it is already in situ and case-marked, but rather as an unscoped wh-phrase in need of a Q-particle) and holds it in working memory until the relevant Q-particle can be located. This process creates an additional working memory load between the wh-element and the Q-particle, similar to the working memory load postulated between a wh-filler and its gap in English or German, and results in sustained anterior negativity. However, this anterior negativity is right- instead of left-lateralized, possibly because participants implicitly imagine the prosodic contour specific to Japanese wh-questions, which in fact mirrors the interrogative scope. At any rate, shared memory resources rather than a special memory system are likely to be used for the processing of filler-gap/wh-Q dependencies, as sustained anterior negativities have also been elicited by sentential stimuli without filler-gap dependencies (Münte et al., 1997; Vos et al., 2001), as well as by non-sentential stimuli (Ruchkin et al., 1990, 1992; Rösler et al., 1993, 1995), as discussed above.

At the scope-marking Q-particle position, the parser associates the wh-phrase and the Q-particle (in other words, completes the wh-Q dependency), thus determining the interrogative scope of the question. However, in contrast to wh-movement languages, this interpretive process does not elicit any additional phasic LAN or P600 effects at the Q-position, most likely because it does not involve either the retrieval or syntactic integration of a displaced wh-filler. Instead, the only necessary operation would be linkage of the Japanese wh-element with its Q-particle in terms of interrogative scope (cf. Nishigauchi, 1990; Cheng, 1991), which does not seem to elicit any discernible local ERP effects.

From the above, one could argue that wh-in-situ questions in Japanese are processed by means of the pre-verbal, incremental formation of a long-distance dependency between wh-elements and their corresponding Q-particles, rather than by a local disambiguation operation that occurs only when the verb-Q position is encountered. If RAN really indexes working memory load as suggested above, the parser would try to remove the extra load of carrying an unresolved wh-Q dependency as soon as possible, possibly by anticipating the site of potential Q-particles. If this view is taken, the results of the present study are then also consistent with previous studies of the processing of Japanese wh-questions (Miyamoto and Takahashi, 2002; Aoshima et al., 2004) which suggest that the parser expects a Q-particle at the earliest grammatically possible position following the appearance of a Japanese wh-element, well before the actual Q-particle position itself.¹⁰ In a broader scheme, this view is not only compatible with a variety of processing models pertaining to filler-gap dependencies, including the first-resort over last-resort strategy (Fodor, 1978; Garnsey et al., 1989), the active filler strategy (Frazier and Clifton, 1989), and the minimal chain principle (de Vincenzi, 1991), in which the parser tries to associate an extracted wh-element with the first possible gap position, but also with incremental

(e.g., Inoue and Fodor, 1995) rather than with head-driven (e.g., Pritchett, 1992) models of language processing.

4.4 Summary and conclusions

This study investigated the processing of Japanese wh-in-situ questions using ERPs. Both mono-clausal and bi-clausal Japanese wh-questions elicited right-lateralized anterior negativity (RAN) between wh-words and corresponding Q-particles, relative to structurally equivalent yes/no-question control conditions. These results suggest a reliable neural processing correlate of the dependency between wh-elements and Q-particles in Japanese, similar to the ERP effects seen between wh-fillers and gaps in English and German, but with a right- rather than left-lateralized distribution. This result along with other behavioral studies (Miyamoto and Takahashi, 2002; Aoshima et al., 2004) suggests that wh-in-situ questions in Japanese are processed by the incremental formation of a long-distance dependency between wh-elements and their corresponding Q-particles, resulting in a working memory load for keeping track of the scopeless wh-elements. The difference between right and left lateralization of the anterior negativity is plausibly due to an implicit prosody effect related to specific pitch contours of Japanese wh-questions, which coincide with the domain of their interrogative scope. Future studies are planned to scrutinize the exact driving forces behind the right-lateralization of this anterior negativity by investigating the processing of different types of syntactic dependencies in Japanese.

5. Experimental procedures

5.1 Experiment 1

5.1.1 Participants—Twenty (11 female) native speakers of Japanese between 19 and 29 years of age (mean: 25) were included in the study.¹¹ They were residents of the San Diego area and had been living outside Japan for less than two years. Participants had normal or corrected-to-normal vision, were right-handed, and had no neurological or reading disorders. Participants were reimbursed for their time.

5.1.2 Materials—Experimental sentences consisted of four versions of 200 mono-clausal questions, namely, (a) wh-in-situ questions, (b) yes/no-in-situ questions, (c) scrambled wh-questions (with the object displaced in front of the subject), and (d) scrambled yes/no-questions, as shown in Table 1 in Section 2. Scrambled conditions were included to increase the distance between wh-elements and Q-particles and to form the basis of another experimental manipulation, as reported in Ueno and Kluender (2003). 200 filler items consisted of four separate question types, namely, (a) ditransitives in canonical word order, (b) ditransitives with accusative-marked demonstrative objects scrambled within the verb phrase, preceding the dative objects, (c) ditransitives with accusative-marked demonstrative objects scrambled within the clause, preceding the nominative subjects, and (d) bi-clausal embedded clause wh-questions with wh-subjects. The 200 sets of sentences instantiating the four experimental conditions were placed in a Latin square design to create four parallel lists of 200 sentences, such that no one participant saw more than one sentence from each set. The 200 filler sentences

¹⁰While Aoshima et al. (2004) reported local slowdown only at points of mismatch, it is not entirely clear whether the reading time slowdown at the embedded verb position of main clause wh-questions in Miyamoto and Takahashi (2002) was due to the claimed local mismatch effect for the wrong particle or more to a continuous dependency between Japanese wh-elements and their corresponding Q-particle as argued in the present study. According to the report of the reading times for the entire sentence positions of the same experiment (Miyamoto and Takahashi, 2003), Miyamoto and Takahashi's main clause wh-questions in fact elicited continuously longer reading times than the other control conditions between embedded verb and main verb positions, showing a parallel with the sustained RAN in response to main clause wh-questions in our ERP data.

¹¹The total number of participants actually run was 27. However, due to experimenter error, four participants saw stimuli at a different presentation rate from the others, so their data were discarded. Three additional participants were excluded, one who exhibited excessive muscle tension that interfered with the EEG recording, and two who never returned for their second session.

were added to each list, and then each list was pseudo-randomized. The materials were divided into 20 sets of 20 sentences each.

5.1.3 Procedure—Participants were run in two sessions lasting about 2.5 hours each. Participants were seated in a comfortable reclining chair facing a computer monitor in a sound-attenuated chamber and wore an elastic cap mounted with tin electrodes. An illuminated rectangular border appeared uninterruptedly in the middle of the screen during presentation of experimental sentences for purposes of fixation. Stimuli were presented in Japanese characters basically one *bunsetsu* at a time with 650 ms duration and 650 ms stimulus onset asynchrony. A *bunsetsu* consists of one free morpheme (lexical word or pronoun) and the bound morpheme (s) associated with it (particles modifying the noun/verb), and is referred to as a “word” elsewhere in the present paper. The presentation rate of 650 ms per word is slower than rates used in many English studies, but was deemed optimal after consulting three native speakers of Japanese. (Given both the visual complexity of Chinese characters often used in Japanese and the fact that many of the Japanese *bunsetsu* translate as multiple English words, it is not surprising that readers needed more time per *bunsetsu*.) There was no interstimulus blank interval between *bunsetsu* for this experiment. The interstimulus interval between sentences was three seconds, and participants were given as much rest as they wished between sets of sentences. In order to engage participants’ attention, they were given a modified probe word recognition task after every five sentences on average but at semi-random intervals, in which they were asked to identify the object of the immediately preceding sentence as either a wh-word or a demonstrative.

5.1.4 Electrophysiological recording—The electroencephalogram (EEG) was recorded from 19 positions, including all standard positions of the international 10/20 system, using tin electrodes mounted in an elastic cap. Reference electrodes were positioned on the two mastoid processes, and the EEG was algebraically referenced off-line to the mean of the electrical activity recorded at these two electrodes. To detect blinks and lateral eye-movements for later correction, additional electrodes were placed beneath the right eye and at the outer canthi of the two eyes. Impedances were kept below 5 K Ω . The EEG was amplified with a bandpass of 0.01 to 100 Hz, digitized at 250 Hz, and stored on hard disk for offline analysis. Data with excessive blinks were corrected using a spatial filter algorithm (Dale, 1994), and a band-pass filter set from 0.2 to 15 Hz was used on all the data prior to running analyses to reduce high frequency noise.

5.1.5 Data analysis—Measurements were taken of single-word averages to look for transient effects, as well as of four-word averages to look for longer-lasting effects. Single-word averages consisted of 1000 ms epochs, including a 100 ms prestimulus baseline, while four-word averages consisted of 2700 ms epochs (4×650 ms SOA plus a 100 ms prestimulus baseline).

The statistical analyses were done separately on midline (Fz, Cz, and Pz), parasagittal (Fp1/2, F3/4, C3/4, P3/4, O1/2), and temporal (F7/8, T3/4, T5/6) electrodes. All analyses consisted of repeated measures ANOVAs run on mean amplitudes of specified latency intervals (P200: 200–300 ms; N400: 300–500 ms; phasic LAN/RAN: 300–600 ms; P600: 500–800 ms) unless otherwise noted. Midline analyses included two within-group factors, with two levels of experimental condition type and three levels of anterior/posterior sites. Parasagittal analyses included three within-group factors, with two levels of condition type, five levels of anterior/posterior sites, and two levels of hemisphere. Temporal analyses included three within-group factors, with two levels of condition type, three levels of anterior/posterior sites, and two levels of hemisphere. In addition, trend analyses were performed to investigate P200 and N350 effects (see Section 2.1.1). An alpha level of .05 was used for all statistical tests, with a p-value up to .10 considered marginally significant. The Huynh-Feldt correction for lack of sphericity was

applied whenever applicable (Huynh and Feldt, 1976). Original degrees of freedom are reported with the corrected probability level.

5.2 Experiment 2

5.2.1 Participants—Twenty (13 female) native speakers of Japanese between 19 and 29 years of age (mean: 23) were included in the study.¹² They were residents of the San Diego area who had been living outside Japan for less than two years, but were a different group of Japanese speakers from those who participated in Experiment 1. Participants had normal or corrected-to-normal vision, were right-handed, had no neurological or reading disorders, and were reimbursed for their time.

5.2.2 Materials—Experimental sentences consisted of four types of bi-clausal questions, namely, (a) embedded clause wh-questions, (b) main clause wh-questions, (c) embedded clause yes/no-questions, and (d) main clause yes/no-questions, as shown in Table 4 in Section 3. Recall that the stimuli in Experiment 1 elicited lexical effects in response to wh-words, namely a larger amplitude P200 related to stroke count, and an earlier and larger N350 component related to word length (Section 2.1.1). This made it impossible to use this sentence position for purposes of baselining subsequent averages in Experiment 1. In order to avoid eliciting these effects in Experiment 2, all wh- and non-wh-modifiers in the embedded object noun phrases were matched for word length. Half of the modifier sets contained wh/non-wh adjective pairs with *donna* (どんな) ‘what.kind.of’ as the wh-modifier, and an adjective such as *atarashii* (新しい) ‘new’ as the non-wh counterpart, with ‘what.kind.of’ always written with *kana* and adjectives written with mixtures of *kana* and *kanji*. The other half contained numeral classifier pairs, such as *nan-bai-no* (何杯の) ‘what-cup-of (how many cups of)’ as the wh-modifier, and *ni-hai-no* (二杯の) ‘2-cup-of’ as the non-wh counterpart, always written with two *kanji* characters (with the character for ‘what’ 何 more visually complex than those for numbers, such as ‘two’ 二) and one *kana* character. The overall design of wh- and non-wh- modifiers for this experiment succeeded in eliminating P200 and N350 effects when all the relevant ERPs were averaged together, but yielded reduced N400 amplitude in response to ‘what.kind.of’ (see fn. 6). Main clause verbs for (a) and (c) (‘say’ verbs) and for (b) and (d) (‘ask’ verbs; see Table 4) were matched for word length and *kanji* ratio. The token frequency (based on Amano and Kondo 2000) of the two groups of the verbs did not differ significantly.

200 sets of experimental sentences containing these four conditions were constructed. 150 filler items consisted of four types of declarative sentences or questions: (a) bi-clausal declaratives, (b) mono-clausal ditransitive declaratives, (c) mono-clausal ditransitive yes/no-questions, and (d) mono-clausal ditransitive wh-questions. The 200 sets of experimental sentences were placed in a Latin square design to create four parallel lists of 200 sentences such that no one participant saw more than one sentence from each set. The 150 filler sentences were added to each list, and then each list was pseudorandomized. The materials were divided into 10 sets of 35 sentences each.

5.2.3 Procedure and data analysis—The experimental procedure and the data analysis methods were the same as in Experiment 1 except for the following: participants were run in only one session lasting about 2.5 hours; stimuli were again presented one word (or *bunsetsu*) at a time with a stimulus onset asynchrony of 650 ms but with a duration of 450 ms and an interstimulus interval of 200 ms; the modified probe recognition task, which presented a verb followed by two nouns and asked participants to choose the subject of the verb in the preceding sentence (adopted from Nagata, 1993; Miyamoto and Takahashi, 2001) was this

¹²The total number of participants actually run was 22. However, one participant showed low comprehension scores, and there was an equipment malfunction during the session of another participant. Thus their data were discarded.

time inserted in the stimuli every seven sentences on average but at semi-random intervals. Following a modified Bonferroni procedure (Keppel, 1982) that allows up to [number of conditions -1] contrasts to maintain the same alpha level, an alpha level of .05 was also used for planned pairwise comparisons.¹³

Acknowledgments

This research was supported by NIHDC02503-01A1; Mieko Ueno was supported in part by NIHDC00041-08/09 through the Center for Research in Language at UCSD, a Joseph Naiman graduate fellowship in Japanese Studies at UCSD, MH19554 through the Department of Psychology at the University of Illinois, and a Faculty Research Award in the Department of Linguistics at the University of Oregon. We would like to thank Kara Federmeier, Janet Dean Fodor, Susan Garnsey, Hiroko Hagiwara, Nayoung Kwon, S.-Y. Kuroda, Marta Kutas, Edson Miyamoto, John Moore, Maria Polinsky, Laurie Stowe, Cyma Van Petten, and three anonymous reviewers for helpful comments and discussion, Edson Miyamoto and Shoichi Takahashi for letting us adopt their stimuli, Yasu-Hiko Tohsaku for help with materials, Naoyuki Osaka and Mariko Osaka for providing their Japanese version of the reading span test, Ronald Ohst, H. Wind Cowles, Marc Silver, and Ezra Van Everbroeck for technical support, and Anne Formanek and Cindy Kilpatrick for editorial support.

References

- Amano, S.; Kondo, T. *Nihongo-no Goi-Tokusei* (Lexical Properties of Japanese). Sanseido, Tokyo.: 2000.
- Aoshima S, Phillips C, Weinberg A. Processing filler-gap dependencies in a head-final language. *J Mem Lang* 2004;51:23–54.
- Bader, M. Prosodic influence on reading syntactically ambiguous sentences. In: Fodor, JD.; Ferreira, F., editors. *Reanalysis in Sentence Processing*. Kluwer Academic Publishers; Dordrecht: 1998. p. 1-46.
- Barrett SE, Rugg MD. Event-related potentials and the semantic matching of faces. *Neuropsychologia* 1989;27:913–922. [PubMed: 2771030]
- Cheng, L. Doctoral Dissertation. MIT; 1991. On the topology of wh-questions.
- Chomsky, N. *Lectures on Government and Binding*. Foris; Dordrecht, Netherlands: 1981.
- Chomsky, N. *Knowledge of Language: Its Nature, Origin, and Use*. Praeger/Greenwood; New York: 1986.
- Chomsky, N. *The Minimalist Program*. MIT Press; Cambridge, MA: 1995.
- Coulson S, King JW, Kutas M. Expect the unexpected: Event-related brain response to morphosyntactic violations. *Lang Cogn Processes* 1998;13:21–58.
- Crain, S.; Fodor, JD. How can grammars help parsers?. In: Dowty, D.; Karttunen, L.; Zwicky, A., editors. *Natural Language Parsing: Psychological, Computational, and Theoretical Perspectives*. Cambridge University Press; Cambridge, UK: 1985. p. 94-128.
- Dale, AM. Doctoral Dissertation. UCSD; 1994. Source localization and spatial discriminant analysis of event-related potentials: Linear approaches.
- Dambacher M, Kliegl R, Hofmann M, Jacobs AM. Frequency and predictability effects on event-related potentials during reading. *Brain Res* 2006;1084:89–103. [PubMed: 16545344]
- Daneman M, Carpenter PA. Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior* 1980;19:450–466.
- De Vincenzi M. Filler-gap dependencies in a null subject language: Referential and nonreferential WHs. *J Psycholinguist Res* 1991;20:197–213.
- Deguchi, M.; Kitagawa, Y. Prosody and wh-questions. In: Hirotani, M., editor. *Proceedings of the Thirty-Second Annual Meeting of the North Eastern Linguistic Society*. Vol. 32. 2002. p. 73-92.

¹³As specific comparisons were preplanned in the present study, we did not run omnibus ANOVAs, a choice justified by statisticians such as G. Keppel:

When “tests designed to shed light on these particular questions are planned before the start of an experiment, ...a researcher is not interested in the omnibus F test—this test is more appropriate in the absence of specific hypotheses...analytical comparisons can be conducted directly on a set of data without reference to the significance or non-significance of the omnibus F test.” (Keppel, 1982: 106)

- Eckstein K, Friederici AD. Late interaction of syntactic and prosodic processes in sentence comprehension as revealed by ERPs. *Cogn Brain Res* 2005;25:130–143.
- Eckstein K, Friederici AD. It's early: Event-related potential evidence for initial interaction of syntax and prosody in speech comprehension. *J Cogn Neurosci* 2006;18:1696–1711. [PubMed: 17014374]
- Fehr T, Code C, Herrmann M. Common brain regions underlying different arithmetic operations as revealed by conjunct fMRI-BOLD activation. *Brain Res* 2007;1172:93–102. [PubMed: 17822681]
- Felser C, Clahsen H, Münte TF. Storage and integration in the processing of filler-gap dependencies: An ERP study of topicalization and wh-movement in German. *Brain Lang* 2003;87:345–354. [PubMed: 14642537]
- Fiebach CJ, Schlesewsky M, Friederici AD. Syntactic working memory and the establishment of filler-gap dependencies: Insights from ERPs and fMRI. *J Psycholinguist Res* 2001;30:321–338. [PubMed: 11523277]
- Fiebach CJ, Schlesewsky M, Friederici AD. Separating syntactic memory costs and syntactic integration costs during parsing: The processing of German WH-questions. *J Mem Lang* 2002;47:250–272.
- Fodor JD. Parsing strategies and constraints on transformations. *Linguistic Inquiry* 1978;9:427–473.
- Fodor JD. Empty categories in sentence processing. *Lang Cogn Processes* 1989;4:155–209.
- Fodor, JD. Prosodic disambiguation in silent reading. In: Hirotani, M., editor. *Proceedings of North East Linguistic Society*. Vol. 32. GLSA, University of Massachusetts; Amherst: 2002. p. 113–132.
- Frazier L, Clifton CJ. Successive cyclicity in the grammar and the parser. *Lang Cogn Processes* 1989;4:93–126.
- Friederici AD, Frisch S. Verb argument structure processing: The role of verb-specific and argument-specific information. *J Mem Lang* 2000;43:476–507.
- Friederici AD, Hahne A, Mecklinger A. Temporal structure of syntactic parsing: Early and late event-related brain potential effects. *J Exp Psychol Learn Mem Cogn* 1996;22:1219–1248. [PubMed: 8805821]
- Fuster, JM. *The Prefrontal Cortex: Anatomy, Physiology, and Neuropsychology of the Frontal Lobe*. Lippincott-Raven Publishers; Philadelphia, PA: 1997.
- Garnsey SM, Tanenhaus MK, Chapman RM. Evoked potentials and the study of sentence comprehension. *J Psycholinguist Res* 1989;18:51–60. [PubMed: 2926696]
- Goldman-Rakic, PS. Circuitry of primate prefrontal cortex and regulation of behavior by representational memory. In: Mountcastle, V., editor. *Handbook of Physiology: The Nervous System*. Vol. 5. American Physiological Society; Bethesda, MD: 1987. p. 373–417.
- Gouvea, A.; Phillips, C.; Kazanina, N.; Poeppel, D. The linguistic processes underlying the P600. submittedMs
- Hagiwara H, Soshi T, Ishihara M, Imanaka K. A topographical study on the event-related potential correlates of scrambled word order in Japanese complex sentences. *J Cogn Neurosci* 2007;19:175–193. [PubMed: 17280508]
- Hagoort P, Brown CM, Groothusen J. The syntactic positive shift (SPS) as an ERP measure of syntactic processing. *Lang Cogn Processes* 1993;8:439–483.
- Hauk O, Pulvermüller F. Effects of word length and frequency on the human event-related potential. *Clinical Neurophysiology* 2004;115:1090–1103. [PubMed: 15066535]
- Hirose, Y. Doctoral Dissertation. CUNY; 1999. Resolving reanalysis ambiguity in Japanese relative clauses.
- Hirotani M. Prosodic effects on the interpretation of Japanese wh-questions. *University of Massachusetts Occasional Papers in Linguistics* 2003;27:117–137.
- Hirotani M, Frazier L, Rayner K. Punctuation and intonation effects on clause and sentence wrap-up: Evidence from eye movements. *J Mem Lang* 2006;54:425–443.
- Hsiao JH, Shillcock R, Lee C. Neural correlates of foveal splitting in reading: Evidence from an ERP study of Chinese character recognition. *Neuropsychologia* 2007;45:1280–1292. [PubMed: 17098263]
- Huynh H, Feldt LS. Estimation of the box correction for degrees of freedom from sample data in randomized block and split-plot designs. *Journal of Educational Statistics* 1976;1:69–82.

- Inoue, A.; Fodor, JD. Information-paced parsing of Japanese. In: Mazuka, R.; Nagata, N., editors. *Japanese Sentence Processing*. Erlbaum, Hillsdale; NJ: 1995. p. 9-63.
- Ishihara, S. Invisible but audible wh-scope marking: Wh-constructions and deaccenting in Japanese. In: Mikkelsen, L.; Potts, C., editors. *Proceedings of the 21st West Coast Conference on Formal Linguistics*. Cascadia Press; Somerville, MA: 2002. p. 180-193.
- Kaan E, Harris A, Gibson E, Holcomb P. The P600 as an index of syntactic integration difficulty. *Lang Cogn Processes* 2000;15:159–201.
- Keppel, G. *Design and Analysis: A Researcher's Handbook*. Englewood Cliffs, NJ: Prentice-Hall; 1982.
- King JW, Kutas M. Who did what and when? using word- and clause-level ERPs to monitor working memory usage in reading. *J Cogn Neurosci* 1995a;7:376–395.
- King JW, Kutas M. A brain potential whose latency indexes the length and frequency of words. *CRL Newsletter* 1995b;10:1–9.
- King JW, Just MA. Individual differences in syntactic processing: The role of working memory. *J Mem Lang* 1991;30:580–602.
- King JW, Kutas M. Neural plasticity in the dynamics of human visual word recognition. *Neurosci Lett* 1998;244:61–64. [PubMed: 9572585]
- Kluender R, Kutas M. Bridging the gap: Evidence from ERPs on the processing of unbound dependencies. *J Cogn Neurosci* 1993a;5:196–214.
- Kluender R, Kutas M. Subjacency as a processing phenomenon. *Lang Cogn Processes* 1993b;8:573–633.
- Kluender, R.; Münte, TF. ERPs to grammatical and ungrammatical wh questions in German: Subject/object asymmetries. Poster Session Presented at the 11th Annual CUNY Conference on Human Sentence Processing.; 1998.
- Koelsch S, Gunter T, Friederici AD, Schröger E. Brain indices of music processing: “Nonmusicians” are musical. *J Cogn Neurosci* 2000;12:520–541. [PubMed: 10931776]
- Koelsch S, Jentschke S. Short-term effects of processing musical syntax: An ERP study. *Brain Res* 2008;1212:55–62. [PubMed: 18439987]
- Kwon, N. Doctoral Dissertation. UCSD; 2008. Processing of syntactic and anaphoric gap-filler dependencies in Korean: Evidence from self-paced reading time, ERP and eye-tracking experiments.
- Liu Y, Perfetti CA, Hart L. ERP evidence for the time course of graphic, phonological, and semantic information in Chinese meaning and pronunciation decisions. *Learning, Memory* 2003;29:1231–1247.
- Liu Y, Perfetti CA, Hart L. ERP evidence for the time course of graphic, phonological, and semantic information in Chinese meaning and pronunciation decisions. *J Exp Psychol Learn Mem Cogn* 2003;29:1231–1247. [PubMed: 14622057]
- Martín-Loeches M, Casado P, Gonzalo R, De Heras L, Fernández-Frías C. Brain potentials to mathematical syntax problems. *Psychophysiology* 2006;43:579–591. [PubMed: 17076814]
- Martín-Loeches M, Muñoz F, Casado P, Melcón A, Fernández-Frías C. Are the anterior negativities to grammatical violations indexing working memory? *Psychophysiology* 2005;42:508–519. [PubMed: 16176373]
- McCallum, WC. Human slow potential research: A review. In: McCallum, WC.; Curry, SH., editors. *Slow Potential Changes in the Human Brain*. Plenum; New York: 1993. p. 1-12.
- McKinnon R, Osterhout L. Constraints on movement phenomena in sentence processing: Evidence from event-related brain potentials. *Lang Cogn Processes* 1996;11:495–523.
- Mecklinger A. On the modularity of recognition memory for object form and spatial location: A topographic ERP analysis. *Neuropsychologia* 1998;36:441–460. [PubMed: 9699951]
- Meyer M, Alter K, Friederici AD, Lohmann G, von Cramon DY. fMRI reveals brain regions mediating slow prosodic modulations in spoken sentences. *Hum Brain Mapp* 2002;17:73–88. [PubMed: 12353242]
- Miyamoto, ET.; Takahashi, S. The processing of wh-phrases and interrogative complementizers in Japanese. In: Akatsuka, N.; Strauss, S., editors. *Japanese/Korean Linguistics*, 10. Vol. 10. CSLI; Stanford: 2002. p. 62-75.
- Miyamoto, ET.; Takahashi, S. Typing mismatch effects in the processing of wh-phrases in Japanese. 2003. Ms

- Molinaro N, Vespignani F, Job R. A deeper reanalysis of a superficial feature: An ERP study on agreement violations. *Brain Res* 2008;1228:161–176. [PubMed: 18619420]
- Müller HM, King JW, Kutas M. Event-related potentials elicited by spoken relative clauses. *Cogn Brain Res* 1997;5:193–203.
- Münter TF, Schiltz K, Kutas M. When temporal terms belie conceptual order. *Nature* 1998;395:71–73. [PubMed: 9738499]
- Münter TF, Wieringa BM, Weyerts H, Szentkuti A, Matzke M, Johannes S. Differences in brain potentials to open and closed class words: Class and frequency effects. *Neuropsychologia* 2001;39:91–102. [PubMed: 11115658]
- Nagata H. Unimmediate construction of syntactic structure for garden path sentences in Japanese. *J Psycholinguist Res* 1993;22:365–381.
- Nakagawa A. Visual and semantic processing in reading kanji. *J Exp Psychol Hum Percept Perform* 1994;20:864–875. [PubMed: 8083640]
- Nakagome K, Takazawa S, Kanno O, Hagiwara H, Nakajima H, Itoh K, Koshida I. A topographical study of ERP correlates of semantic and syntactic violations in the Japanese language using the multichannel EEG system. *Psychophysiology* 2001;38:304–315. [PubMed: 11347875]
- Nakamura K, Oga T, Okada T, Sadato N, Takayama Y, Wydell T, Yonekura Y, Fukuyama H. Hemispheric asymmetry emerges at distinct parts of the occipitotemporal cortex for objects, logograms and phonograms: A functional MRI study. *Neuroimage* 2005;28:521–528. [PubMed: 16027011]
- Neville HJ, Mills DL, Lawson DS. Fractionating language: Different neural subsystems with different sensitive periods. *Cerebral Cortex* 1992;2:244–258. [PubMed: 1511223]
- Neville HJ, Nicol JL, Barss A, Forster KI, Garrett MF. Syntactically based sentence processing classes: Evidence from event-related brain potentials. *J Cogn Neurosci* 1991;3:151–165.
- Nishigauchi, T. *Quantification in the Theory of Grammar*. Kluwer Academic Publishers; Dordrecht, Netherlands: 1990.
- Osterhout L, Nicol J. On the distinctiveness, independence, and time course of the brain responses to syntactic and semantic anomalies. *Lang Cogn Processes* 1999;14:283–317.
- Osterhout L, Bersick M, McKinnon R. Brain potentials elicited by words: Word length and frequency predict the latency of an early negativity. *Biol Psychol* 1997;46:143–168. [PubMed: 9288411]
- Osterhout L, Holcomb PJ. Event-related brain potentials elicited by syntactic anomaly. *J Mem Lang* 1992;31:785–806.
- Osterhout L, Holcomb PJ, Swinney DA. Brain potentials elicited by garden-path sentences: Evidence of the application of verb information during parsing. *J Exp Psychol Learn Mem Cogn* 1994;20:786–803. [PubMed: 8064247]
- Patel AD, Gibson E, Ratner J, Besson M, Holcomb PJ. Processing syntactic relations in language and music: An event-related potential study. *J Cogn Neurosci* 1998;10:717–733. [PubMed: 9831740]
- Phillips C, Kazanina N, Abada SH. ERP effects of the processing of syntactic long-distance dependencies. *Cogn Brain Res* 2005;22:407–428.
- Pritchett, BL. *Grammatical Competence and Parsing Performance*. University Of Chicago Press; Chicago: 1992.
- Rayner K, Kambe G, Duffy SA. The effect of clause wrap-up on eye movements during reading. *The Quarterly Journal of Experimental Psychology Section A* 2000;53:1061–1080.
- Rösler F, Heil M, Glowalla U. Monitoring retrieval from long-term memory by slow event-related brain potentials. *Psychophysiology* 1993;30:170–182. [PubMed: 8434080]
- Rösler F, Heil M, Hennighausen E. Exploring memory functions by means of brain electrical topography: A review. *Brain Topogr* 1995;7:301–313. [PubMed: 7577328]
- Ruchkin DS, Johnson R Jr, Canoune H, Ritter W. Short-term memory storage and retention: An event-related brain potential study. *Electroencephalogr Clin Neurophysiol* 1990;76:419–439. [PubMed: 1699736]
- Ruchkin DS, Johnson R Jr, Canoune H, Ritter W. Event-related potentials during arithmetic and mental rotation. *Electroencephalogr Clin Neurophysiol* 1991;79:473–487. [PubMed: 1721575]

- Ruchkin DS, Johnson R Jr, Grafman J, Canoune H, Ritter W. Distinctions and similarities among working memory processes: An event-related potential study. *Cogn Brain Res* 1992;1:53–66.
- Saito M. Some asymmetries in Japanese and their theoretical implications. 1985
- Saito M. Long distance scrambling in Japanese. *Journal of East Asian Linguistics* 1992;1:69–118.
- Stowe LA. Parsing WH-constructions: Evidence for on-line gap location. *Lang Cogn Processes* 1986;1:227–245.
- Takashima A, Ohta K, Shirahama Y, Matsushima E, Horiguchi T, Nishikawa T. Difference of semantic processing and character types in silent reading of words—a study using MEG 2002;1232:513–517.
- Takashima A, Ohta K, Matsushima E, Toru M. The event-related potentials elicited by content and function words during the reading of sentences by patients with schizophrenia. *Psychiatry Clin Neurosci* 2001;55:611–618. [PubMed: 11737794]
- Ueno, M. Doctoral Dissertation. UCSD; 2003. Event-related brain potentials in the processing of Japanese wh-questions.
- Ueno M, Garnsey SM. An ERP study of the processing of subject and object relative clauses in Japanese. *Lang Cogn Processes* 2008;23:646–688.
- Ueno M, Kluender R. Event-related brain indices of Japanese scrambling. *Brain Lang* 2003;86:243–271. [PubMed: 12921767]
- Van Petten C, Kutas M. Interactions between sentence context and word frequency in event-related brain potentials. *Mem Cognit* 1990;18:380–393.
- Van Petten C, Kutas M. Influences of semantic and syntactic context in open- and closed-class words. *Mem Cognit* 1991;19:95–112.
- Van Petten C, Kutas M, Kluender R, Mitchiner M, McIsaac H. Fractionating the word repetition effect with event-related potentials. *J Cogn Neurosci* 1991;3:131–150.
- Vos SH, Gunter TC, Kolk HH, Mulder G. Working memory constraints on syntactic processing: An electrophysiological investigation. *Psychophysiology* 2001;38:41–63. [PubMed: 11321620]
- Yamaguchi S, Toyoda G, Xu J, Kobayashi S, Henik A. Electroencephalographic activity in a flanker interference task using Japanese orthography. *J Cogn Neurosci* 2002;14:971–979. [PubMed: 12419122]
- Yamashita H. Scrambled sentences in Japanese: Linguistic properties and motivations for production. *Text* 2002;22:597–633.
- Yamashita H. The effects of word-order and case marking information on the processing of Japanese. *J Psycholinguist Res* 1997;26:163–188.
- Ye Z, Luo Y, Friederici AD, Zhou X. Semantic and syntactic processing in Chinese sentence comprehension: Evidence from event-related potentials. *Brain Res* 2006;1071:186–196. [PubMed: 16412999]
- Yoshida H, Iwahashi M, Yamaguchi M, Tonoike M. Neuromagnetic activity in the human left cerebral hemisphere concerning logical processing during auditory oddball stimulation. *Neuroreport* 1998;9:1261–1265. [PubMed: 9631409]
- Zatorre RJ, Evans AC, Meyer E, Gjedde A. Lateralization of phonetic and pitch discrimination in speech processing. *Science* 1992;256:846–849. [PubMed: 1589767]

Figure 1.

(a) ERPs ($n = 20$) for all electrodes to wh-words and demonstratives, collapsed across sentence positions; (b) ERPs ($n = 20$) at midline electrode Pz to four types of pronominal elements, collapsed across sentence positions.

Figure 2. ERPs (n = 20) for all electrodes relative to a 100 ms prestimulus baseline. Shown are the ERPs to the final verb-Q position (‘discovered-Q’ in Table 1) of wh-in-situ vs. yes/no-in-situ-questions. The isovoltage map is based on the mean difference calculated as the wh-in-situ minus yes/no-in-situ conditions for the 300–600 ms interval.

Figure 3. ERPs ($n = 20$) for all electrodes relative to a 100 ms prestimulus baseline following wh-words up to and including the verb-question particle complex. Shown are the ERPs covering the adjective, subject noun, adverbial, and verb-question-particle positions ('reckless adventurer-NOM finally discovered-Q' in Table 1) of scrambled wh- vs. yes/no-questions. The isovoltage map is based on the mean difference calculated as the scrambled wh minus scrambled yes/no conditions for the 950–2600 ms, 950–1250 ms, 1600–1900 ms, and 2250–2550 intervals.

Figure 4.

ERPs ($n = 20$) at right anterior electrode F8 relative to a 100 ms prestimulus baseline. Shown are the ERPs to the subject noun ('reckless adventurer-NOM'), adverbial ('finally'), and verb-question-particle ('discovered-Q') positions of scrambled wh- vs. yes/no-questions.

..... All Wh									
専務が	どんな	パソコンを	買ったか/と	経理の	係長が	聞き/言いましたか。			
[Director-NOM	what.kind.of	computer-ACC	bought-Q/that]	accounting-of	manager-NOM	__ asked/said-Q.			
_____ All Yes/No									
専務が	新しい	パソコンを	買ったか/と	経理の	係長が	聞き/言いましたか。			
[Director-NOM	new	computer-ACC	bought-Q/that]	accounting-of	manager-NOM	__ asked/said-Q.			

Figure 5. ERPs (n=20) recorded from all electrodes relative to a 100 ms prestimulus baseline. Shown are the ERPs to the embedded object noun phrase position ('what.kind.of/new computer-ACC' in Tables 4 and 5) of all wh- vs. yes/no-questions. The isovoltage maps are based on the mean difference calculated as all wh minus all yes/no conditions for the 500–1250 ms, 500–700 ms, and 950–1250 ms intervals.

Figure 6.

ERPs ($n=20$) recorded from all electrodes relative to a 100 ms prestimulus baseline. Shown are the ERPs to the embedded verb position ('bought-Q/that') of main clause wh, embedded clause wh, and main clause yes/no conditions. The isovoltage maps are based on the mean difference calculated as the main-wh minus main-yes/no conditions (above) and the main-wh minus embedded-wh conditions (below) for the 300–900 ms interval.

Figure 7.

ERPs ($n=20$) recorded from all electrodes relative to a 100 ms prestimulus baseline. Shown are the ERPs to the region from the embedded to the main clause verb position ('bought-Q/ that accounting-of manager-NOM asked-Q' in Tables 4 and 5) of main clause wh, embedded clause wh, and main clause yes/no conditions. The isovoltage maps are based on the mean difference calculated as the main-wh minus main-yes/no conditions (above) and the main-wh minus embedded-wh conditions (below) for the 300–2600 ms, 300–600 ms, 950–1250 ms, 1600–1900 ms, and 2250–2450 ms intervals.

Figure 8.

ERPs ($n=20$) at right anterior electrode F8 relative to a 100 ms prestimulus baseline. Shown are the ERPs to the embedded verb ('bought-Q/that'), noun modifier ('accounting-of'), main clause subject noun ('manager-NOM'), and main clause verb ('asked-Q') positions of main clause wh, main clause yes/no, and embedded clause wh conditions.

Table 1

Sample stimuli for Experiment 1

a. Object wh-in-situ questions (wh-in-situ questions)

あの 地元の 新聞に よると

Ano jimotono shinbun-ni yoruto
the local newspaper-to according

その 命知らずの 冒険家が とうとう 何を 見つけたんですか。

sono inochishirazuno bokenka-ga toto **nani-o** mitsuketa-ndesu-ka.
the/that reckless adventurer-NOM finally what-ACC discovered- POL- Q

'According to the local newspaper, what did that reckless adventurer finally discover?'

b. Yes/no-questions with demonstrative objects in situ (yes/no-in-situ questions)

あの 地元の 新聞に よると

Ano jimotono shinbun-ni yoruto
the local newspaper-to according

その 命知らずの 冒険家が とうとう それを 見つけたんですか。

sono inochishirazuno bokenka-ga toto **sore-o** mitsuketa-ndesu-ka.
the/that reckless adventurer-NOM finally that-ACC discovered- POL- Q

'According to the local newspaper, did that reckless adventurer finally discover that?'

c. Scrambled object wh-questions (scrambled wh-questions)

あの 地元の 新聞に よると

Ano jimotono shinbun-ni yoruto
the local newspaper-to according

何を その 命知らずの 冒険家が とうとう 見つけたんですか。

nani-o sono inochishirazuno bokenka-ga toto — mitsuketa-ndesu-ka.
what-ACC the/that reckless adventurer-NOM finally discovered- POL- Q

'According to the local newspaper, what did that reckless adventurer finally discover?'

d. Yes/no-questions with scrambled demonstrative objects (scrambled yes/no- questions)

あの 地元の 新聞に よると

Ano jimotono shinbun-ni yoruto
the local newspaper-to according

それを その 命知らずの 冒険家が とうとう 見つけたんですか。

sore-o sono inochishirazuno bokenka-ga toto — mitsuketa-ndesu-ka.
that-ACC the/that reckless adventurer-NOM finally discovered- POL- Q

'According to the local newspaper, did that reckless adventurer finally discover that?'

Table 2

F-values for ANOVAs for wh vs. non-wh word comparisons in Experiment 1

Analysis	Effect (df)	P200	N350	
		Mean Amp. 200–300 ms	Peak Latency 250–450 ms	Mean Amp. 300–375 ms
Midline	C (1,19)	4.28 *	5.67 **	4.30 *
	C x A (2,38)	ns	ns	ns
Parasagittal	C (1,19)	4.01 *	3.86 *	3.64 *
	C x H (1,19)	5.64 **	ns	6.90 **
	C x A (4,76)	ns	2.37 *	ns
	C x H x A (4,76)	ns	ns	ns
Temporal	C (1,19)	3.43 *	6.84 **	5.66 **
	C x H (1,19)	20.03 ***	ns	11.58 ***
	C x A (2,38)	7.87 ***	ns	ns
	C x H x A (2,38)	ns	ns	ns

C: Condition, A: Anteriority, H: Hemisphere

*
p < .10**
p < .05***
p < .01

Table 3
F-values for ANOVAs for scrambled wh vs. yes/no question comparisons in Experiment 1

Analysis	Effect (df)	adventurer-N				discovered-O	
		950–2600 ms	950–1250 ms	1600–1900 ms	2250–2550 ms		
Midline	C (1,19)	ns	ns	ns	ns		
	C x A (2,38)	ns	ns	ns	3.11*		
	C (1,19)	ns	ns	ns	ns		
Parasagittal	C x H (1,19)	**	**	**	6.72**		
	C x A (4,76)	ns	ns	ns	ns		
	C x H x A (4,76)	ns	ns	ns	ns		
Temporal	C (1,19)	ns	ns	ns	3.57*		
	C x H (1,19)	ns	ns	ns	ns		
	C x A (2,38)	ns	ns	ns	ns		
	C x H x A (2,38)	**	**	2.99*	ns		

* p < .10
** p < .05
*** p < .01

Table 4

Sample stimuli for Experiment 2

a. Embedded clause wh-questions

専務が どんな パソコンを買ったか 経理の 係長が 聞きましたか。
 [senmu-ga donna pasokon-o katta-ka] keiri-no kakaricho-ga kikimashita-ka.
 director-N what.kind.of computer-A bought-Q accounting-of manager-N asked.POL-Q

'Did the accounting manager ask what kind of computer the director bought?'

b. Main clause wh-questions

専務が どんな パソコンを買ったと 経理の 係長が 言いましたか。
 [senmu-ga donna pasokon-o katta-to] keiri-no kakaricho-ga iimashita-ka.
 director-N what.kind.of computer-A bought-that accounting-of manager-N said.POL-Q

'what kind of computer did the accounting manager say the director bought?'

c. Embedded clause yes/no-questions

専務が 新しい パソコンを買ったか 経理の 係長が 聞きましたか。
 [senmu-ga atarashii pasokon-o katta-ka] keiri-no kakaricho-ga kikimashita-ka.
 director-N new computer-A bought-Q accounting-of manager-N asked.POL-Q

'Did the accounting manager ask if the director bought a new computer?'

d. Main clause yes/no-questions

専務が 新しい パソコンを買ったと 経理の 係長が 言いましたか。
 [senmu-ga atarashii pasokon-o katta-to] keiri-no kakaricho-ga iimashita-ka.
 director-N new computer-A bought-that accounting-of manager-N said.POL-Q

'Did the accounting manager say the director bought a new computer?'

Table 5

Points of comparisons and ERP effects

a. All wh

'Did the accounting manager ask what kind of computer the director bought?'
'What kind of computer did the accounting manager say the director bought?'
'Did the accounting manager ask if the director bought a new computer?'
'Did the accounting manager say that the director bought a new computer?'

Table 6

F values for ANOVAs for main clause wh vs. main clause yes/no and main clause wh vs. embedded clause wh comparisons in Experiment 2

Analysis	Effect (df)	Main clause wh vs. Main clause yes/no						Main clause wh vs. Embedded clause wh					
		bought-that			accounting-of			bought-that			Accounting-of		
		300-2600ms	300-1950ms	950-1250ms	1600-1900ms	2250-2550ms	said-Q	300-2600ms	300-600ms	950-1250ms	1600-1900ms	2250-2550ms	said-Q
Midline	C (1,19)	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns
	C x A (2,3)	ns	3.87**	ns	ns	2.98*	2.57*	2.98*	ns	ns	10.42***	ns	ns
Prasagittal	C (1,19)	ns	ns	ns	ns	ns	ns	3.20*	ns	ns	2.88*	5.25**	ns
	C x H (1,1)	ns	ns	3.16*	ns	ns	ns	ns	ns	ns	ns	ns	ns
	C x A (4,7)	ns	ns	4.25**	ns	ns	ns	ns	ns	ns	ns	4.38**	ns
	C x H x A	ns	ns	ns	ns	2.75*	ns	ns	ns	ns	ns	ns	ns
Temporal	C (1,19)	ns	ns	ns	ns	ns	ns	4.71**	ns	ns	ns	9.74***	ns
	C x H (1,1)	ns	4.70***	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns
	C x A (2,3)	ns	2.51*	ns	2.73*	ns	ns	ns	ns	ns	ns	ns	ns
	C x H x A	ns	ns	ns	ns	ns	ns	ns	ns	2.96*	ns	ns	ns

C: Condition, A: Anteriority, H: Hemisphere

* p < .10

** p < .05

*** p < .01