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

Mieko Ueno
Department of Psychology and Beckman Institute for Advanced Science and Technology,
University of Illinois, Urbana-Champaign

Robert Kluender
Department of Linguistics, University of California, San Diego

Running head: ERPS TO JAPANESE WH-QUESTIONS

June 2004

Address correspondence to:
Mieko Ueno
2217 Beckman Institute
University of Illinois, Urbana-Champaign
405 N. Matthews
Urbana, IL 61801

phone: (217)244-1952
fax: (217)333-8206
email: ueno@psych.uiuc.edu
Abstract

An event-related brain potential (ERP) experiment investigated the processing of Japanese wh-questions. Unlike 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 if there are 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 the scope of a Japanese wh-element is licensed by a long-distance incremental linkage with its Q-particle, rather than by a local scope calculation process at the Q-particle.
Acknowledgements

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 dissertation fellowship in the Department of Linguistics at UCSD, a Joseph Naimam graduate fellowship in Japanese Studies at UCSD, and MH19554 through the Department of Psychology at UIUC. The authors would like to thank Kara Federmeier, Janet Dean Fodor, Susan Garnsey, Hiroko Hagiwara, S.-Y. Kuroda, Marta Kutas, Reiko Mazuka, Edson Miyamoto, John Moore, Maria Polinsky, Laurie Stowe, and Cyma Van Petten for helpful comments and discussion, Amalia Arvaniti and Mary Beckman for help with the prosody analysis, Yasu-Hiko Tohsaku for help with materials, Ronald Ohst, H. Wind Cowles, and Sean Perry for technical support, and Anne Sumnicht for editorial support.
On the Processing of Japanese Wh-Questions: An ERP Study

Introduction

Japanese: A Wh-in-situ Language

Wh-questions—i.e., questions that contain interrogative 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; Stowe 1986; Frazier & Clifton 1989, among others). In terms of electrophysiological processes, there have been a number of studies that have investigated the processing of wh-questions using event-related brain potentials (ERPs) (Kluender & Kutas 1993a,b; McKinnon and Osterhout 1996; Kluender & Münte 1998; Kaan, Harris, Gibson, & Holcomb (2000); Fiebach, Schlesewsky, & Friederici, 2001, Phillips, Kazanina, Wong, & Ellis, 2001; Phillips, Kazanina, Garcia-Pedrosa, & Abada, 2003; Felser, Clahsen, & Münte, 2003).

These studies have been done in English and German, so-called "wh-movement" languages that require wh-elements to be displaced to the beginning of a clause. Consider (1a): the wh-object *what* is displaced to the beginning of the clause compared to the non-wh object *pizza* in (1b).

(1)  a. What did Calvin bring __?
    b. Did Calvin bring pizza?

It has been proposed that associating the displaced wh-element (*what* in (1a)) with its canonical position (indicated by underlining in (1a), and occupied by *pizza* in (1b)) increases verbal working memory load, and that this processing cost is reflected in an ERP component known as left anterior negativity (LAN).
Unlike English and German, Japanese is a “wh-in-situ” language where wh-words stay in the same canonical SOV (subject-object-verb) position as their non-wh counterparts. As shown in (2), ‘pizza’ and ‘what’ stay in the same position.

(2)

a. Calvin-ga ピザを 持ってきたんですか。 Calvin-NOM(INATIVE) pizza-ACC(USATIVE) brought-POLITE-Q(UESTION) ‘Did Calvin bring pizza?’

b. Calvin-ga 何を 持ってきたんですか。 Calvin-NOM nani-o motteka-ndesu-ka. ‘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 and 1992 for syntactic considerations and Yamashita 1997 and 2002 for processing considerations).

While displacement 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).² Wh-questions without a Q-particle are ungrammatical in Japanese, as shown in (3b).

(3)

a. Calvin-ga 何を 持ってきたんですか。 Calvin-NOM nani-o motteka-ndesu-ka. ‘What did Calvin bring?’

b. *Calvin-ga 何を 持ってきたんですよ。 *Calvin-NOM nani-o motteka-ndesu-yo. ‘Calvin brought what.’

In addition, this 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. The
surface position of moved wh-elements in wh-movement languages, such as English and German, transparently indicates their interrogative clausal scope relations in 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”), or at the beginning of the main clause as in (4b), yielding a main clause wh-question (also termed a “direct question”). Thus, 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. On the other hand, the logical answer to a matrix clause wh-question like (4b) would be the referent of the wh-word (pizza in this case), although in daily conversation one might actually provide the referent of the wh-word in answer for both questions for pragmatic reasons.

(4) Wh-Movement Languages (e.g., English)

a. Embedded clause wh-question

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

b. Main clause wh-question

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

In Japanese, instead of the surface position of a wh-element, the position of its related Q-particle indicates its interrogative scope. As shown in (5), while the wh-elements in the embedded clause wh-question (5a) and the main clause wh-questions (5b) stay in the embedded clause, the Q-particle ka appears at the end of the interrogative clause, either embedded (5a) or main (5b).
(5) a. Embedded clause wh-question

\[ \text{Hobbes-wa Calvin-ga nani-o mottekita-ka itta-ndesu-ka]? } \]

Hobbes-TOPIC Calvin-NOM what-ACC brought-Q said-POLITE-Q

‘Did Hobbes say what Calvin brought?’

b. Main clause wh-question

\[ \text{Hobbes-wa Calvin-ga nani-o mottekita-to itta-ndesu-ka]? } \]

Hobbes-TOPIC Calvin-NOM what-ACC brought-that said-POLITE-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 as shown in (6), a wh-element is displaced to the left of its gap and marks its scope in English, while in Japanese a Q-particle is placed to the right of the corresponding wh-element and marks its scope.

(6) a. English

\[ \text{[What did Hobbes say Calvin brought __ ]]? } \]

FILLER GAP

b. Japanese

\[ \text{Hobbes-wa Calvin-ga nani-o mottekita-to itta-ndesu-ka]? } \]

Hobbes-TOPIC Calvin-NOM what-ACC brought-that said-POLITE-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.

**Processing of Japanese Wh-questions**

Given the above, one might expect processing effects specific to wh-in-situ. One possible effect would be related to the scope ambiguity of wh-elements, i.e., whether a certain wh-element should be interpreted as a main clause interrogative or an embedded clause interrogative. 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 surface position of wh-elements, the parser in a wh-in-situ language like Japanese will be unable to determine the clausal scope of a wh-element until the relevant related question particle is encountered. Recall that in (5) both sentences are identical up to the particle (to, 'that', or ka, 'Q') position. Thus it is possible that the parser has to hold a wh-element in working memory until the corresponding particle position before it can determine its scope within the sentence, causing an extra processing load.

There has been experimental evidence suggesting some type of 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 and without Q-particles at the sentence-final verb position. Wh-questions without Q-particles elicited effects of late positivity. Nakagome et al. attributed this so-called P600 effect to syntactic processes related to movement operations.

In a self-paced reading experiment, Miyamoto and Takahashi (2002) reported that main clause wh-questions were read more slowly than embedded clause wh-questions at the embedded verb position when there was no Q-particle attached to the embedded verb. Miyamoto and Takahashi argued that just like the dependency 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. Main clause wh-questions were read more slowly than embedded clause wh-questions because they violated this expectation (Typing Mismatch Effect: TME). Similarly, Aoshima, Phillips, and Weinberg (in press) observed a TME slowdown even when a wh-element placed in the main clause had to be interpreted as scrambled out of the embedded clause.

Predictions

The present study examines the processing of mono-clausal and bi-clausal Japanese wh-questions. The aim of the study was to examine the extent to which the neural processing of wh-questions in a wh-in-situ language, such as Japanese, shows similarities to the processing of wh-questions in wh-movement languages, such as English and German. More specifically, based on the above experimental studies that suggest some type of relationship between a wh-element and its corresponding Q-particle, we hypothesized that, similar to filler-gap dependencies in English or German, there might be a dependency between Japanese wh-elements and their Q-particles. Further, we predicted that if there really is a dependency between a Japanese wh-element and its Q-particle *ka*, and if the wh-Q dependency is similar to a filler-gap dependency, we may see the same type of ERP components, such as LAN and slow negative anterior potentials (Kluender & Kutas, 1993a,b; King & Kutas, 1995; Kluender & Münte, 1998; Fiebach et al., 2001), between Japanese wh-words and related Q-particles, as shown in (7).
(7) Wh-Q dependency hypothesis

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

```
What did Calvin bring ___?
FILLER                     GAP
   (L)AN
```

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

```
カルビンが  何を  持ってきたんですか。
Calvin-NOM  nani-o  mottekita-ndesu-ka
Calvin-NOM  what-ACC  brought- POLITE-Q
   WH                     Q
   (L)AN?
```

‘What did Calvin bring?’

In addition, as a wh-element occurring in situ needs its interrogative scope licensed by some mechanism, we hypothesized that there might be an interpretive process involved in calculating the scope of such wh-element. We further hypothesized that this “scope calculation effect” might be seen at the Q-particle position that determines the scope of the wh-element, as shown in (8).

(8) Local scope calculation hypothesis

```
scope-calculation effect?
LAN?, P600??
```

```
カルビンが  何を  持ってきたんですか。
Calvin-ga  nani-o  mottekita-ndesu-ka
Calvin-NOM  what-ACC  brought- POLITE-Q
   WH                     Q
   (L)AN?
```

‘What did Calvin bring?’

However, as this was to our knowledge the first ERP study of grammatical wh-in-situ processing, it was difficult to predict specific ERP component(s) that might index such a process. If the parser has to recall the words contained in the interrogative clause for purposes of scope
calculation, thereby taxing working memory load, one might reasonably expect an effect of left anterior negativity at the sentence-final verb-Q position. If one assumes with Kaan et al. (2000) and Phillips et al. (2001, 2003) that the P600 is an index of syntactic integration in sentence-processing contexts, or with Felser et al. (2003) that the P600 is an index of an operator-variable dependency, one might reasonably expect an effect of late positivity at the verb-Q position. On the other hand, given that Japanese is a left-branching, head-final, wh-in-situ language, the possibility also existed that the processing of wh-in-situ questions in Japanese could elicit some other type of specific ERP effect.

Experiment 1

Experiment 1 investigated the processing of mono-clausal Japanese wh-questions as a first step.

Methods

Subjects.

Twenty (11 female) monolingual speakers of Japanese between 19 and 29 years of age (mean: 25) who had been outside of Japan for less than two years were included in the study. Subjects had normal or corrected-to-normal vision, were right-handed, and had no neurological or reading disorders. Subjects were reimbursed for their time.

Materials.

Stimuli consisted of four conditions of mono-clausal questions, namely, (a) wh-in-situ questions, (b) yes/no-in-situ questions, (c) scrambled wh-questions, and (d) scrambled yes/no-questions, as shown in Table 1. Scrambled conditions were included to increase the distance between wh-elements and Q-particles. Filler items consisted of four separate types of interrogative sentences, namely, (a) ditransitives in canonical word order, (b) ditransitives with
accusative-marked demonstrative objects scrambled within the VP, preceding the dative objects, (c) ditransitives with accusative-marked demonstrative objects scrambled within S (IP), preceding the nominative subjects, and (d) bi-clausal embedded clause wh-questions with wh-subjects. 200 sets of stimulus sentences containing the four experimental conditions were constructed and placed in a Latin square design to create four parallel lists of 200 stimulus sentences, such that no one subject saw more than one sentence from each set. The 200 filler sentences were added to each list, and then each list was pseudo-randomized. The materials were divided into 20 sets of 20 sentences each.

---Table 1---

*Procedure.*

Subjects were run in two sessions lasting about 2.5 hours each. Subjects were seated in a reclining chair facing a computer monitor in a sound-attenuated room 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 on a computer screen in Japanese characters basically one *bunsetsu* at a time with 650 ms duration and 650 ms⁵ 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 will be referred to as a “word” hereafter. Due to experimenter error, there was no interstimulus interval between each *bunsetsu*⁶ for this experiment. The interstimulus interval between sentences was three seconds, and subjects were given as much rest as they wished between sets of sentences. In order to maintain subjects' attention, comprehension questions were inserted in the stimuli. Every five sentences on average
but at a semi-random interval, subjects were asked to answer a comprehension question regarding the immediately preceding sentence.

*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 activity 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 5KΩ. The EEG was amplified with a bandpass of 0.01 to 100 Hz, digitized at 250 Hz, and stored on hard disk for off-line 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.

*Data analysis.*

Measurements were taken of both single-word averages for phasic effects, and four-word averages mid-sentence in the scrambled wh and scrambled yes/no conditions for longer-lasting effects. Single-word averages consisted of 1024 ms epochs, including a 100 ms prestimulus baseline, while four-word averages consisted of 3072 ms epochs (4 x 650 ms SOA plus a 400 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. Midline analyses consisted of repeated measures ANOVAs with two within-group factors, including two levels of experimental condition type and three levels of anterior/posterior sites. Parasagittal analyses
ERPs to Japanese Wh-Questions 14

consisted of repeated measures ANOVAs with three within-group factors, including two levels of condition type, five levels of anterior/posterior sites, and two levels of hemisphere. Temporal analyses consisted of repeated measures ANOVAs with three within-group factors, including 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 2.2). An alpha level of .05 was used for all statistical tests, with a p-value of .10 considered marginally significant. The Huynh-Feld correction for lack of sphericity was applied whenever applicable. Original degrees of freedom are reported with the corrected probability level.

Results

The mean correct response rate to comprehension questions across subjects was 91% (S.D. 5.2%). Thus no subject's data were excluded from the ERP analyses based on poor comprehension. Sentences were compared at wh-elements to see if there was any effect of wh-processing at wh-elements themselves, and from the words following wh-elements to the sentence-final verb-Q complex to see if there were any effects following localized lexical processing of wh-elements.

Single-word averages of wh- and demonstrative pronouns.

As a first step, ERPs (collapsed across sentence positions) to two types of wh-pronouns (‘what-ACC’ 何を and ‘who-ACC’ 誰を) and two types of demonstrative pronouns (‘that-ACC’ を and ‘that-person-ACC’ あのを) used in the experiment were compared to see if there was any effect of wh-processing at wh-words themselves. Visual inspection of waveforms suggested larger P200 amplitude in response to wh-words relative to demonstratives. In addition, the negative peak at around 350 ms (N350) following the P200 to wh-words was larger and earlier than that to demonstratives. These differences in the P200 and N350 between wh-
words and demonstratives initially gave the impression of a wh-processing-effect. However, a closer examination of four different pronominal elements (see Figure 1) suggested that this was merely an effect of lexical processing differences.

Regarding the P200 effect, it was noticed that wh-pronouns include a higher ratio of graphically complex Chinese characters requiring more strokes to write. Trend analyses on P200 amplitude, collapsed over electrodes for the four pronouns, revealed a significant linear component for total stroke count \( F(1, 57) = 14.21, p < .001 \), but none of the other higher-order (quadratic, cubic) trend components were significant. Correlations between P200 amplitude and frequency or word length did not reveal a significant linear trend component.

As for the N350 effect, in light of claims that correlate the negative peak latency following the P200 with frequency and word length in lexical processing paradigms (Neville, Mills, & Lawson, 1992; Osterhout, Bersick, & McKinnon, 1997; King & Kutas, 1998), trend analyses were used to determine if N350 peak latency, collapsed over electrodes for the four pronouns, was influenced by frequency or word length. While frequency did not, word length did reveal a significant linear trend component \( F(1, 57) = 6.89, p = 0.012 \) but no significant higher trend components. As further possibilities, the orthogonal trends between N350 latency and stroke count were tested, yet revealed no significant linear trend. For the amplitude difference of the N350, mean amplitudes collapsed across electrodes for the four pronouns were tested with the possible factors examined above. Among stroke count, frequency, and word length, word length was again the only factor that revealed a significant linear trend component \( F(1, 57) = 4.76, p = .033 \) by itself.
Between wh and verb-Q: wh vs. yes/no-questions.

As a second step following the wh-word comparison, ERPs to words following wh-words up to and including the verb-Q (question particle) complex were examined. The two in-situ conditions and two scrambled conditions ((a) vs. (b) and (c) vs. (d) in Table 1) were compared separately, as it was assumed that comparing conditions identical in word order would reveal effects of wh-processing alone.

Figure 2 shows the comparison between wh-in-situ questions vs. in-situ demonstrative yes/no-questions at the sentence-final verb (‘discovered-Q’ in Table 1 a, b). Wh-in-situ questions appeared to elicit greater negativity after 300 ms post-stimulus onset of ‘discovered-Q’, especially over the right side of the head. This observation was in part statistically supported by ANOVAs performed within the latency window of 300-600 ms. The temporal array showed a significant main effect \[F(1, 19) = 4.69, p = .043\] and a marginal condition x hemisphere interaction \[F(1, 19) = 3.90, p = .063\] that indicated a greater right-lateralized negativity in response to wh-in-situ questions compared to yes/no-in-situ questions. In addition, although not statistically significant, the negativity to wh-in-situ questions appeared to be somewhat posterior in its distribution.

Figure 3 shows the comparison of words following scrambled wh- and demonstrative pronouns up to and including the verb-Q complex (‘reckless adventurer-NOM finally discovered-Q’ in Table 1 c, d). Visual inspection of these four-word averages showed a right-lateralized negativity at about 500 ms following the onset of ‘finally’ to the sentence end of the scrambled wh-question condition.
To capture the first observation, ANOVAs were performed in the latency window of 500 ms poststimulus onset of ‘finally’ to the sentence end, including the sentence final verb-Q complex ‘discovered-Q’ (1800 to 2600 ms from the baseline). A significant main effect of condition was found in all three arrays [midline: $F(1, 19) = 6.66, p = .018$; parasagittal: $F(1, 19) = 6.19, p = .022$; temporal: $F(1, 19) = 5.26, p = .033$]. In addition, there was a marginal interaction between condition and anteriority in the midline array [$F(2, 38) = 3.13, p = .088$], as well as a significant interaction between condition and hemisphere in the parasagittal array [$F(1, 19) = 8.78, p < .001$]. These interactions were due to greater negativity over right anterior regions in response to scrambled wh-sentences.

To summarize, scrambled wh-questions in comparison to scrambled yes/no-questions elicited greater negativity, especially over the right side of the head, both at the adverbial and at the verb-Q complex positions ending the stimulus sentences.

Summary of effects.

Wh-words elicited larger P200s and a subsequently delayed, larger amplitude negative peak in the latency range of 250 to 450 ms (N350) compared to demonstratives. The P200 correlated with stroke count, whereas both the latency and the amplitude of the N350 correlated solely with word length. Wh-in-situ questions compared to yes/no-in-situ questions elicited greater right-lateralized negativity at the sentence-final verb-Q position, while scrambled wh-questions elicited greater right-lateralized negativity than scrambled yes/no-questions at both the penultimate adverbial and sentence-final verb-Q positions.
Discussion

Lexical effects on wh-words.

Wh-words in comparison to demonstratives elicited larger P200s and later and larger N350s, with P200 amplitude correlating with stroke count, and N350 amplitude/latency with word length. The interaction between P200 amplitude and stroke count most likely reflects local visual processing of complex Chinese characters, consistent with reports of larger P200s when local features are attended to (Kotchoubey, Wascher, & Verleger, 1997) and claims relating the P200 to visual targets and attention (Luck & Hillyard, 1994; O’Donnell, Swearer, Lloyed, Hokama, 1997). Additionally, Takashima, Ohta, Matsushima, and Toru (2001) reported that content words (nouns, verbs) elicit larger P200s than function words (particles, auxiliary verbs). However, this effect is confounded by the fact that content words tend to be written in Chinese characters, while function words are usually written in visually less complex phonemic characters. Our results, showing that even graphically complex closed-class words elicit larger P200s, strongly suggest that visual complexity rather than word class is the crucial variable in P200 amplitude. As for the N350, word length was the only factor that showed a significant linear trend for both the later peak latency and the larger amplitude. This goes along with Osterhout et al.’s (1997) claim, in that longer demonstratives elicited later peaks than shorter wh-words. As for the peak amplitude difference, Neville et al. (1992) reported that among open class words, longer words elicited more positive ERPs than shorter words following the P200 over frontal regions (Neville et al. 1992: Figure 5). Although the morphology of their data differs from that of our data, in both cases, longer words (demonstratives in the present study) elicited more positive (or less negative) ERPs than shorter words (wh-words in the present study) after the P200. It is interesting that frequency, which has been reported to exert a strong
influence over lexical ERP effects in English (Osterhout et al., 1997; King & Kutas, 1998), played no apparent role in the present study. We initially presumed that since wh-words and demonstratives were both closed-class words, they would provide a reasonable basis for investigating wh-processing effects. However, the comparison revealed only lexical processing effects.

*Effect of wh between wh and verb-Q complex.*

Although the right-lateralized negativity occurring between wh-words and the verb-Q complex in both in-situ and scrambled comparisons was not exactly identical in latency and anteriority, there were several aspects in common. One was question type: both effects occurred in response to wh-questions relative to non-wh (yes/no)-questions. Others include: point of occurrence (after processing wh-words), polarity (negative), and lateralization (right-lateralized). From the above, it seems plausible to consider the two effects at least 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 effect. One is as an end-of-sentence wrap-up effect. However, previous reports of such effects involve N400-like negativities with posterior distribution at the end of sentences containing syntactic and semantic anomalies (Osterhout & Holcomb, 1992; Hagoort, Brown, & Groothusen, 1993; Osterhout & Nicol, 1999) or a widely distributed negativity also at the end of sentences containing syntactic and semantic anomalies (Friederici & Frisch, 2000). Unlike studies with anomalies, the stimulus sentences used in the present experiment were all grammatical, and the distribution of the negativity was anterior at least for the scrambled wh- vs. yes/no-question comparison. These points diminish the possibility of a sentence-end wrap-up effect.
Another possibility is that the sentence-ending negativity indexed a dispreferred parse effect. However, dispreferred parse effects in otherwise grammatical sentences have been linked to the P600 around the critical word (Osterhout & Holcomb, 1992; Osterhout, Holcomb, & Swinney, 1994; Friederici, Hahne, & Mecklinger, 1996) rather than to any sentence-end negativity to the best of our knowledge. For instance, among the sentence types with dispreferred constructions, sentence-end negativity was found only in response to ungrammatical conditions (Osterhout et al., 1994: Figure 5). In other words, grammatical dispreferred parses seem to result in P600 effects at sentence-intermediate positions and no apparent sentence-final negativities, rendering it unlikely that the right-lateralized negativity effects in the present study could be taken as an index of a dispreferred parse.

One could also speculate that subjects may have been imagining different intonation contours for wh- and yes/no-questions at sentence-end (the critical word in the present paradigm), causing different brain responses. The possibility of an implicit prosody effect was examined in a production study of sample stimuli for both mono-clausal and bi-clausal wh-questions in Experiment 3. To foreshadow those results, it is unlikely that implicit prosody played a significant causal role in the present study.

In view of the unlikelihood of these other possibilities, the right-lateralized effect most likely indexes some type of relationship between Japanese wh-words and Q-particles. This is consistent with previous experimental claims about Japanese wh-questions, as discussed in the Introduction. Thus, our preliminary conclusion at this point is that the brain is sensitive to the relationship between a wh-word and a following Q-particle. However, since the effect did not appear to be a slow potential (the negativity did not start until the final two words of the sentence for the scrambled comparison), there are two possible interpretations. One is that the sentence-
ending negativity is an index of the parser’s expectation for a Q-particle, (“wh-Q expectation hypothesis”). The other is that the negativity indexes a more local integration of wh-scope at the Q-particle position that disambiguates the scope of the wh-element (“local scope calculation hypothesis”). Experiment 2 investigated the processing of bi-clausal Japanese wh-questions to tease apart these two possibilities.

Experiment 2

Recall from Experiment 1 that mono-clausal Japanese wh-questions elicited right-lateralized negativity (anterior for the scrambled comparison) negativity around the sentence-final position. This indicated some type of relationship between wh-words and Q-particles, rather than a sentence-end wrap-up, dispreferred parse, or implicit prosody effect (see Experiment 3). The remaining question was what this right (anterior) negativity effect might be indexing in terms of language processing with regard to the “wh-Q dependency” and “local scope calculation” hypotheses. The aim of Experiment 2 was to address this question using bi-clausal wh-questions involving embedded and main clause scope, as these were expected to involve various types of scope calculation-related operations.

The stimuli 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\textsuperscript{10} as shown in (9) with our predictions.
(9) Stimuli for Experiment 2 and Predictions

a. Embedded clause wh-question

\[
\begin{align*}
\text{wh-Q RAN?} & \quad \text{local scope RAN?} \\
\downarrow & \\
[\text{senmu-ga donna pasokon-o katta-ka}] & \quad [\text{keiri-no kakaricho-ga kikimashita-ka}]
\end{align*}
\]

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

b. Main clause wh-question

\[
\begin{align*}
\text{wh-Q RAN?} & \quad \text{local scope RAN?} \\
\downarrow & \\
[\text{senmu-ga donna pasokon-o katta-to}] & \quad [\text{keiri-no kakaricho-ga iiimashita-ka}]
\end{align*}
\]

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

If the wh-Q dependency hypothesis is correct and right (anterior) negativity (hereafter referred to as “RAN”) indexes the parser’s on-going expectation for a Q-particle, it should manifest between the wh-element and embedded verb-Q positions in embedded clause wh-questions (9a), and between the wh-element and the main clause verb-Q in main clause wh-questions (9b). If on the other hand the local scope calculation hypothesis is correct and RAN indexes the parser’s local calculation of wh-scope at the sentence position where wh-scope becomes disambiguated, then it should manifest at the embedded verb-Q position of embedded clause wh-questions (9a), and at the main clause verb-Q position of main clause wh-questions (9b).
Methods

Subjects.

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

Materials.

Stimuli consisted of four conditions of bi-clausal questions, namely, (a) embedded clause wh-questions, (b) matrix clause wh-questions, (c) embedded clause yes/no-questions, and (d) matrix clause yes/no-questions, as shown in Table 2. Recall that the stimuli in Experiment 1 elicited lexical effects in response to wh-words, namely larger amplitude P200 and N350 components correlated with stroke count and word length. In order to see if there was any effect of wh-processing beyond these purely lexical factors, all of the wh- and non-wh-modifiers in the embedded object noun phrases were matched for word length. As shown in Table 3, half of the stimulus sets contained 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. Each pair had equal numbers of characters, but different Chinese character (kanji) to phonemic character (kana) ratios (e.g., *donna* どんな ‘what.kind.of’ consists of all kana, while *atarashii* 新しい ‘new’ consists of one kanji and two kana segments). The other half of the stimulus sets contained numeral classifier pairs, such as *nan-bai-no* (何杯の) ‘what-cup-of (how many cups of)’ as the wh-modifier, and numeral classifiers such as *ni-hai-no* (二杯の) ‘2-cup-of’ as the non-wh counterpart. These wh- and non-wh-classifiers were written in essentially the same way
except for one character, *nan* ([₀]&f1; ‘what’ vs. *ni* ‘2’ ([₀]), which would minimize stroke count effects. Main clause verbs for (a) and (c) (‘say’ verbs) and for (b) and (d) (‘ask’ verbs) were matched for character count and Chinese character (*kanji*) ratio. The token frequency of the two groups of the verbs did not differ significantly.12

**Table 2**

**Table 3**

200 sets of sentences containing these four conditions were constructed. Filler items consisted of four conditions of declarative or interrogative sentences, namely, (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 stimulus sentences were placed in a Latin square design to create four parallel lists of 200 stimulus sentences such that no one subject 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.

**Procedure.**

The procedure was the same as in Experiment 1 except for the following. Subjects were run in one session lasting about 2.5 hours each. Stimuli were presented one word (or *bunsetsu*) at a time with 650 ms duration and 450 ms stimulus onset asynchrony. Comprehension questions were inserted in the stimuli every seven sentences on average but at a semi-random interval. In addition to single- and four-word averages, two-word averages were computed with 2048 ms epochs including a 200 ms prestimulus baseline. 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
Results

The mean correct response rate to comprehension questions across subjects was 87% (S.D. 8.0%). Comparisons were made to examine the two hypotheses discussed in the Introduction (i.e., the wh-Q dependency hypothesis and the local scope calculation hypothesis).

Recall that the local scope calculation hypothesis predicts that if RAN indexes the parser’s local calculation of wh-scope at the sentence position where wh-scope becomes disambiguated, it should be seen at the embedded verb-Q position in embedded clause wh-questions, and at the main clause verb-Q position in main clause wh-questions. This hypothesis was tested by measuring ERPs at the embedded and main clause verb positions. By contrast, the wh-Q dependency hypothesis predicts that if RAN indexes subjects’ expectation for a Q-particle, it should be seen between wh-elements and Q-particles in both embedded and main clause wh-questions. This hypothesis was tested by investigating ERPs between wh-elements and Q-particles.

Wh-modifier.

Recall that in 50% of the stimulus sets, the wh-modifier of the embedded object noun (‘what.kind.of PC-ACC’ in Table 3) was donna ‘what.kind.of’; its non-wh-counterpart was an adjective such as atarashii ‘new’ (‘new PC-ACC’ in Table 3). The other 50% of the stimulus sets contained numeral classifiers such as nan-bai-no ‘what-cup-of (how many cups of)’ in the wh-conditions and ni-hai-no ‘2-cup-of’ in the non-wh-condition. ERPs to both groups of wh-modifiers were compared to their non-wh counterparts in order to examine if there was an effect of Q-expectancy at the position of the wh-modifier itself.

Figure 4 shows ERPs to ‘what.kind.of’ and adjective modifiers such as atarashii ‘new’. An effect of anterior negativity in response to ‘what.kind.of’, as well as P200 and N400 effects
ERPs to Japanese Wh-Questions

in response to adjectives were observed. In order to corroborate the effect of anterior negativity, ANOVAs were run in the latency window of 300-900 ms\textsuperscript{14}. There was a significant or marginal condition x anteriority interaction in all arrays [\textit{midline}: $F(2, 38) = 3.1, p = .09$; \textit{parasagittal}: $F(4, 76) = 4.49, p = .018$; \textit{temporal}: $F(2, 38) = 8.8, p = .001$], indicating that ‘what.kind.of’ was more negative than adjectives over anterior regions.

As for the P200 effect to adjectives, ANOVAs run in the 200-300ms window revealed a significant or marginal main effect of condition in all three arrays [\textit{midline}: $F(1, 19) = 5.29, p = .033$; \textit{parasagittal}: $F(1, 19) = 5.03, p = .037$; \textit{temporal}: $F(1, 19) = 3.86, p = .064$], as well as a significant condition x anteriority interaction in the parasagittal [$F(4, 76) = 3.77, p = .031$] and temporal [$F(2, 38) = 8.99, p = .005$] arrays. These results indicated that adjectives had a larger P200 than ‘what.kind.of’, especially over anterior regions. As for the N400 effect to adjectives, separate ANOVAs were run in the window of 350-450 ms. There was a significant or marginal main effect of condition [\textit{midline}: $F(1, 19) = 10.34, p = .005$; \textit{parasagittal}: $F(1, 19) = 4.85, p = .04$; \textit{temporal}: $F(1, 19) = 3.79, p = .067$], as well as a significant or marginal condition x anteriority interaction [\textit{midline}: $F(2, 38) = 4.2, p = .052$; \textit{parasagittal}: $F(4, 76) = 8.61, p = .001$; \textit{temporal}: $F(2, 38) = 15.17, p = .001$] in all three arrays. These results indicated that the N400 to adjectives was larger than that to ‘what.kind.of’, especially over posterior sites.

--Figure 4--

Figure 5 shows ERPs to the numeral classifier pairs, such as ‘what-cup-of’ vs. ‘2-cup-of’. Increased right-lateralized negativity was observed in response to the wh-condition. ANOVAs run in the window of 300-600 ms showed a significant main effect of condition in the parasagittal array [$F(1, 19) = 5.24, p = .034$] and a marginal condition x hemisphere interaction in the temporal array [$F(1, 19) = 4.03, p = .059$], indicating that the ERPs to the ‘what-cup-of’
condition were more negative, especially over the right hemisphere. Additionally, ANOVAs run
in the window of 300-900 ms showed a significant main effect of condition in all three arrays
[\textit{midline}: F(1, 19) = 4.5, p = .047; \textit{parasagittal}: F(1, 19) = 7.68, p = .012; \textit{temporal}: F(1, 19) =
6.14, p = .023], again indicating that the ERPs to the ‘what-cup-of’ condition were more negative
than the ERPs to the ‘2-cup-of’ condition. Furthermore, a small P200 effect to wh-modifiers in
comparison to non-wh-modifiers (this time the opposite of the ‘what.kind.of’ vs. ‘new’
comparison that elicited a larger P200 in response to non-wh-modifiers) was observed, yet
ANOVAs run in the 200-300ms window revealed no significant or marginal main effect.

--Figure 5--

To summarize, ERPs to wh-modifiers revealed RAN effects in comparison to non-wh
modifiers. The comparison between \textit{donna} ‘what.kind.of’ and adjectives revealed anterior
negativity in response to wh-modifiers, as well as larger amplitude P200 and N400 components
in response to adjectives. The comparison between the numeral classifier pairs revealed only
right-lateralized negativity in response to wh-modifiers.

\textit{Embedded object region.}

The glosses of the relevant experimental sentences with the point of comparison
capitalized are given below in (10).
Both types of (i.e., both embedded clause and main clause) wh-questions and yes/no-questions were collapsed, as they were identical up to the embedded verb position.

Figure 6 shows ERPs to all wh-questions in comparison to all yes/no-questions at the embedded object position. Visual inspection showed a sustained anterior negativity in response to all wh-questions. ANOVAs run in the window of 300-1300ms revealed a significant or marginal main effect of condition in all three arrays [midline: F(1, 19) = 3.28, p = .086; parasagittal: F(1, 19) = 5.83, p = .026; temporal: F(1, 19) = 4.81, p = .041], as well as a marginal condition x anteriority interaction in the temporal array [F(2, 38) = 3.75, p = .063]. These results indicated that wh-questions were more negative than yes-no questions, especially over fronto-central regions of scalp.

---Figure 6---

To summarize, similar to the wh-modifier position, an effect of anterior negativity was found in response to the embedded object position of wh-questions. This indicates that the negativity started at the wh-modifier position, and continued through the embedded object. This effect of anterior negativity cannot be an effect of local wh-scope calculation (i.e. in support of
the local scope calculation hypothesis), as at this point in the sentence, subjects had not seen any Q-particle that would disambiguate the scope of the wh-element they had just seen. On the other hand, an effect of this type is entirely consistent with the wh-Q expectation hypothesis, in so far as the introduction of a wh-word triggers the parser's expectation for a Q-particle.

*Embedded verb.*

The glosses of the relevant experimental sentences with the point of comparison capitalized are given below in (11).

(11)

(a. Embedded clause wh

[Director-N what.kind.of computer-A BOUGHT-Q] accounting-of manager-N asked-Q

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

(b. Main clause wh

[Director-N what.kind.of computer-A BOUGHT-THAT] accounting-of manager-N said-Q

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

(c. Embedded clause yes/no

[Director-N what.kind.of computer-A BOUGHT-Q] accounting-of manager-N asked-Q

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

The two hypotheses make different predictions at this position. If RAN indexes the parser’s expectation for a Q-particle, the effect should be elicited by main clause wh-questions in comparison to embedded clause wh-questions or main clause yes/no-questions, as there should be an unresolved dependency at the embedded verb-particle complex of main clause wh-questions. On the other hand, if RAN indexes an effect of wh-scope calculation, only the embedded clause wh-condition (11a) should elicit a RAN effect, as it is the only condition that requires wh-scope calculation at the scope-disambiguating Q-particle of the embedded clause.
By contrast, scope calculation for the main clause wh-condition (11b) should not occur until the main clause verb-Q position, and there should be no wh-scope calculation for the yes/no control condition (11c) at all.

Figure 7 shows ERPs to embedded vs. main clause wh-questions. Increased RAN was seen in response to main clause wh-questions rather than to embedded clause wh-questions. ANOVAs performed in the 300-600 ms window showed no significant or marginal differences, but ANOVAs in the 300-900 ms window revealed a marginal condition x anteriority x hemisphere interaction in the temporal array \[F(2, 38) = 2.88, p = .071\], indicating that main clause wh-questions in comparison to embedded clause wh-questions were more negative especially over right anterior regions of scalp. Although the effect was statistically weak, this was consistent with the prediction made by the wh-Q expectation hypothesis and contrary to the local wh-scope calculation hypothesis. Although there was a visible N400-like effect to embedded clause wh-questions at posterior electrodes, ANOVAs run in the window of 350-450 ms revealed no significant or marginal differences.

--Figure 7--

Figure 8 shows ERPs to main clause wh- vs. yes/no-questions at the embedded verb ‘bought-that’ position. Similar to the main vs. embedded clause wh-question comparison, a RAN effect was observed in response to main clause wh-questions. ANOVAs performed in the 300-600 ms window showed no significant or marginal differences, while ANOVAs in the 300-900 ms window revealed a significant condition x anteriority interaction in the midline array \[F(2, 38) = 3.74, p = .035\], as well as a marginal condition x hemisphere interaction in the parasagittal \[F(1, 19) = 3.36, p = .083\] and temporal \[F(1, 19) = 3.67, p = .071\] arrays. These results indicate that main clause wh-questions were more negative over right anterior regions.\(^{15}\)
ERPs to embedded clause wh- vs. yes/no-questions were also compared at the embedded verb position. Visual inspection indicated no RAN differences between embedded clause wh- and yes/no-questions. ANOVAs performed in the 300-600 ms and 300-900 ms windows showed no significant or marginal effects. On visual inspection there also appeared to be an N400-like effect at this sentence position—this time, however, in response to embedded yes/no-questions relative to embedded wh-questions. Yet ANOVAs run in the window of 350-450 ms revealed no significant or marginal differences.

To summarize, contrary to predictions of the local scope calculation hypothesis, there were no differences in RAN between embedded and main clause wh-questions at the embedded verb position. Instead, main clause wh-questions elicited a (statistically marginally) RAN effect. Comparisons of embedded clause yes/no-questions with embedded wh-questions revealed no effects. Thus the pattern of results up to this sentence position is in line with the wh-Q dependency hypothesis.

*From embedded to main clause verbs.*

Let us now investigate the latter half of the stimulus sentences, namely from embedded to main clause verbs, and test whether the prediction of the wh-Q dependency hypothesis also holds in this region. The glosses of the relevant experimental sentences with the point of comparison capitalized are given below in (12).

(12)

a. Main clause wh

[Director-N what.kind.of computer-A BOUGHT-THAT] ACCOUNTING-OF MANAGER-N SAID-Q

‘What kind of computer did the accounting manager say the director bought?’
b. Embedded clause wh

[Director-N what.kind.of computer-A BOUGHT-Q] ACCOUNTING-OF MANAGER-N ASKED-Q

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

c. Main clause yes/no

[Director-N what.kind.of computer-A BOUGHT-THAT] ACCOUNTING-OF MANAGER-N SAID-Q

‘Did the accounting manager say that the director bought a new computer?’

In this sentence region, only the main clause wh-question condition (12a) is expected to elicit a RAN effect according to the wh-Q dependency hypothesis, as it is the only condition that still has an unresolved expectation for a Q-particle. The embedded clause wh-question condition (12b) has presumably already fulfilled this expectation at the embedded verb-Q position, and there should be no expectation at all for a Q-particle in the yes/no control condition (12c). Thus comparisons were made between main vs. embedded clause wh-questions and between main clause wh- vs. yes/no-questions in a latency window that covered four words (‘bought-Q/that accounting-of manager-N asked/said-Q’) from the embedded verb to the sentence-final main clause verb. We extended the latency window to the end of the sentence in order to look for slow potential effects between a wh-word and its associated Q-particle that would lend further support to the wh-Q expectation hypothesis.

Figure 9 shows ERPs to embedded vs. main clause wh-questions from embedded to main clause verbs, as in ‘bought-Q/that accounting-of manager-N asked/said-Q’. A sustained anterior negativity was observed. ANOVAs run in the 300-2600 ms window confirmed the observation. There was a significant main effect of condition in the parasagittal [F(1, 19) = 4.95, p = .038] and temporal [F(1, 19) = 5.41, p = .031] arrays, as well as a significant or marginal condition x anteriority interaction in the parasagittal [F(4, 76) = 3.72, p = .066] and temporal [F(2, 38) = 4.44,
These results indicated that main clause wh-questions were more negative than embedded clause wh-questions, especially over anterior regions of scalp.

Figure 10 shows ERPs to matrix clause wh- vs. yes/no-questions between embedded and main clause verbs, as in ‘bought-that accounting-of manager-N said-Q’. On a visual inspection, a sustained anterior negativity was observed in response to main clause wh-questions at prefrontal electrodes. However, ANOVAs run in the 300-2600 ms window revealed no significant or marginal differences.

In summary, main clause wh-questions in comparison to embedded clause wh-questions elicited anterior negativity in the four-word window from the embedded verb to the sentence-final main clause verb position. This result remains consistent with the wh-Q expectation hypothesis. With respect to main clause wh-questions in comparison to main clause yes/no-questions, there was a visual trend of anterior negativity continuing through to the end of the sentence, although it was not statistically supported. We suspect that the lack of a statistically significant difference might be due to the N400-like effect at the sentence end of main clause yes/no-questions. As briefly discussed in the following section addressing the main clause verb, we suspect that subjects did not expect a Q-particle at the end of a sentence that would otherwise be declarative (see Discussion).

Main clause verb.

Although the results thus far overwhelmingly support the wh-Q dependency hypothesis over the local scope calculation hypothesis, comparisons were nevertheless made at the sentence-final main clause verb in order to confirm that RAN is not tied to local scope calculation at this
sentence position. The glosses of relevant experimental sentences with the point of comparison capitalized are given below in (13).

(13)

a. Main clause wh

[Director-N what.kind.of computer-A bought-that] accounting-of manager-N SAID-Q

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

b. Embedded clause wh

[Director-N what.kind.of computer-A bought-Q] accounting-of manager-N ASKED-Q

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

c. Main clause yes/no

[Director-N what.kind.of computer-A bought-that] accounting-of manager-N SAID-Q

‘Did the accounting manager say that the director bought a new computer?’

At this position, according to the local scope calculation hypothesis, only the main clause wh-condition (13a) should elicit a RAN effect, as it was the only condition that required wh-scope calculation at the sentence-final scope-disambiguating Q-particle. It was presumed that scope calculation for the embedded clause wh-condition (13b) should already have been completed at the embedded verb-Q position, and that no wh-scope calculation at all should be required for the yes/no control condition (13c). Thus comparisons were made between embedded vs. main clause wh-questions and between main clause wh- vs. yes/no-questions.

Comparisons of ERPs to main clause and embedded clause wh-questions revealed no observable RAN effect, and ANOVAs performed in the 300-600 ms and 300-900 ms windows showed no significant or marginal differences between the two conditions. However, increased negativity around 300-400 ms was observed in response to main clause wh-questions. ANOVAs on mean amplitude in the 300-400 ms window revealed a significant condition x anteriority
interaction in the midline array \[F(2, 38) = 4.79, p = .039\], indicating that main clause wh-questions were more negative at anterior regions.

Comparisons of ERPs to main clause wh- vs. yes/no-questions revealed no observable RAN effect either, and ANOVAs performed in the 300-900 ms window showed no significant or marginal differences between the two conditions. ANOVAs run in the time window of 300-600 ms showed a marginal condition x hemisphere x anteriority interaction in the parasagittal array \[F(4, 76) = 2.12, p = .089\]. However, this was because main clause yes/no-questions were more negative than main clause wh-questions especially at left central sites. More generally, on visual inspection, main clause yes/no-questions showed an N400-like effect. ANOVAs run in the latency window of 350-450 ms showed a marginal main effect of condition in the midline \[F(1, 19) = 3.60, p = .073\] and parasagittal \[F(1, 19) = 3.09, p = .095\] arrays, indicating a larger N400 peak for main clause yes/no-questions compared to main clause wh-questions. Additionally, there was a significant condition x hemisphere interaction in parasagittal \[F(1, 19) = 5.56, p = .029\] and temporal \[F(1, 19) = 5.50, p = .03\] arrays, as well as a significant condition x hemisphere x anteriority interaction in the parasagittal array \[F(4, 76) = 4.28, p = .009\]. These interactions showed that main clause yes/no-questions were more negative than main clause wh-questions at left central regions. However, although the effect was left-lateralized, the absolute values of the negative peak in response to main clause yes/no-questions were right-lateralized, following the standard distribution of the N400. The difference seemed have been due to greater left-right asymmetry in the main clause wh-question condition.

In summary, similar to the comparisons at the embedded verb-Q position, no comparison at the main clause verb-Q position revealed RAN effects as predicted by the local scope calculation hypothesis. The negative peak in response to main clause wh-questions compared to
embedded clause wh-questions was larger over anterior regions, but this difference had a very
short latency window (within 100 ms), unlike the other RAN effects, which lasted more than 300
ms. For the main clause wh- vs. yes/no comparison, there was an N400-like effect in response to
main clause yes/no-questions. We suspect that this was related to the presence of an unexpected
Q-particle at the end of a sentence that could otherwise have been declarative; recall that filler
material consisted of 75 declarative sentences (see Discussion for more details). Overall, neither
comparisons at the embedded verb position nor comparisons at the main clause verb position
supported the local scope calculation hypothesis.

Discussion

**RAN: local scope calculation or wh-Q dependency?**

To summarize, we have seen effects of anterior negativity in response to wh-questions in
multiple comparisons. The effects found in the present experiment (and the mono-clausal
experiment discussed in Experiment 1) were right-lateralized in the majority of cases, although
not always, and at times statistically marginal. In the following subsections, we discuss their
possible implications.

The results of the present experiment support the wh-Q dependency hypothesis over the
local scope calculation hypothesis. First, there was no local RAN at sentence positions where
wh-scope was disambiguated. At the embedded verb position, although with marginal
significance, main clause wh-questions in comparison to embedded clause wh-questions elicited
a RAN effect. This is consistent with what the wh-Q dependency hypothesis would predict (as a
part of a continuous expectation for a Q-particle between wh and Q), while it is the exact
opposite of what the local scope calculation hypothesis would predict (local scope calculation at
the embedded verb-Q position). At the main clause verb position, there was a larger negative
peak in response to main clause wh-questions relative to embedded clause wh-questions. However, the effect was different from the RAN effects seen in other comparisons in terms of latency (i.e., only in the 300-400ms latency window) and distribution (significant only in the midline array). Thus there was no RAN effect at the sentence-final verb. Secondly, both types of wh-questions elicited RAN effects from the wh-modifier position to the corresponding Q-particle position (as shown in Figures 4-6 for all wh-questions in the embedded clause region and in Figures 7-10 for main clause wh-questions), as predicted by the wh-Q dependency hypothesis.

Therefore, the question raised at the end of Experiment 1 seems to have the following answer: the wh-Q dependency hypothesis has been supported over the local scope calculation hypothesis. Nonetheless, two differences between the results of the mono-clausal and bi-clausal experiments are worth mentioning.

First, in the present bi-clausal experiment there was increased anterior negativity at the wh-modifier position, while there was no anterior negativity seen in response to the wh-pronoun position in the mono-clausal experiment. Instead, only lexical effects influencing P200 amplitude and N350 amplitude and latency were found. However, this might have been because word length was not controlled for between wh-pronouns and demonstratives in the mono-clausal experiment. Recall that the wh-pronouns used in the mono-clausal experiment were shorter than the demonstratives, and that there was a later N350 peak to the longer demonstratives. This N350 peak in response to demonstratives could conceivably have masked the anterior negativity in response to wh-words.

Secondly, in the present experiment the negativity to wh-questions was longer lasting, while in the mono-clausal experiment the negativity was seen only near sentence end. This cannot be attributed to differences in the distance between a wh-element and its corresponding
Q-particle across the two experiments, as the average number of words between wh and Q in the present experiment was 2.4, while that in the mono-clausal experiment was 3. Instead, it may have something to do with the bi-clausal structure of the stimulus materials in the present study. In the present experiment, 73% of wh-questions (100 out of 137 wh-questions) were biclausal; there was only one filler condition with mono-clausal wh-questions. On the other hand, in Experiment 1, only 20% of wh-questions (one filler condition: 25 out of 125 wh-questions) were biclausal. Thus in Experiment 2, when subjects saw wh-elements, they may have immediately expected a bi-clausal structure, and this may have caused an earlier RAN effect, as subjects may have anticipated that some of the wh-Q dependencies (main clause wh-questions in particular) were going to be long.

*Other ERP effects.*

Besides the RAN effect, we have seen several other effects in the present experiment. At the wh-modifier position, there were P200 and N400 effects for adjectives in comparison to *donna* ‘what.kind.of’ (see Figure 4). The P200 effect seems consistent with the finding discussed in Experiment 1 that graphic complexity induces a larger P200. While *donna* ‘what.kind.of’ is written all in phonemic characters (*kana*) as どんな, 97.5% of the adjectives included Chinese characters (*kanji*) that are graphically more complex and that require more strokes to write. The N400 effect to open class adjectives compared to closed class ‘what.kind.of’ is also consistent with previous findings regarding the relationship between the N400 and lexical frequency. A corpus count of *Asahi Shinbun* (Amano & Kondo 2000), a popular Japanese newspaper, shows that ‘what.kind.of’ has a much higher frequency (27,715) than the mean frequency of the adjectives used (6,497). Thus the higher lexical frequency of ‘what.kind.of’ (Van Petten & Kutas, 1991) and its repeated presentation in the experiment (Van
Petten, Kutas, Kluender, Mitchiner, & McIsaac, 1991) would have attenuated the N400 in response to ‘what.kind.of’ but enhanced the N400 in response to adjectives. From the above, it seems plausible to conclude that both effects are lexical. When graphic complexity and open vs. closed class differences in lexical frequency between wh- vs. non-wh-modifiers were more closely controlled in the numeral classifier pairs, as in ‘what-cup-of’ and ‘2-cup-of’, both effects disappeared (Figure 5).

The N400-like effect in response to main clause yes/no-questions at the main clause verb-Q position is unlikely to be due to the word length or frequency of ‘say’ and ‘ask’ verbs used for the two conditions, as each pair was matched for word length, and there was no statistically significant difference in the frequency of the two types of verbs (see Methods). Instead, this is likely to be an effect of seeing an unexpected Q-particle at the end of 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 experimental materials (275 questions (79%) out of the 350 experimental material sentences), subjects might still have expected declaratives by default, by virtue of their life experience. In particular, one filler condition had bi-clausal declaratives, as in ‘The director of the museum said that the millionaire donated many pictures.’ Therefore, a Q-particle at the sentence-final matrix verb position was probably less expected, and this may have caused the N400-like effect. The effect had a right central distribution in absolute values, a typical distribution for the N400.
At the sentence-final main clause verb position, main clause wh-questions in comparison to embedded wh-questions showed a brief negative peak in the window of 300-400 ms, with a significant condition x anteriority interaction at the midline (Figure 7). From the distribution and latency, we have concluded that this is different from the RAN effect seen in other comparisons. This does not appear to be an N400 effect either, due to its anterior distribution and rather short latency. Unlike the Q-particle at the end of main clause yes/no-questions, subjects must have been satisfied to see the Q-particle they had been expecting at the end of main clause wh-questions. From this perspective as well, the effect seems unlikely to be an expectation-related N400 effect. We speculate here that this may be a special effect of wh-scope integration across a clause boundary. We did not see this type of effect at the sentence-final verb-Q position of mono-clausal wh-questions, or at the embedded verb-Q position of embedded wh-questions. Thus we assume that this would not be an effect of wh-scope calculation in general. However, it may be plausible to think that there is an extra processing cost to integrating wh-scope across a clause boundary (see Kluender & Kutas, 1993b and Kluender, 1998 for a discussion of a processing cost for holding and integrating a wh-filler across a clause boundary). Future research is needed to clarify this issue.

*RAN: other possibilities?*

Although there was a concern about a sentence-end wrap-up effect for Experiment 1, this was clearly not the case for the present experiment. There were (R)AN effects not only at sentence end, but also in response to constituents of the embedded clause. These effects started at the wh-modifier position, five words before the final word of the sentence. Aside from the sentence-end wrap-up effect, recall that we considered the possibility of a tacit prosody difference effect for mono-clausal questions in Experiment 1. In order to explore this possibility
for Experiment 2 as well, a prosody analysis of sample tokens from Experiments 1 and 2 was conducted in Experiment 3.

Summary of Experiment 2.

Experiment 2 investigated the processing of bi-clausal Japanese wh-questions. There was no clear local RAN effect at the embedded verb-Q position in embedded clause wh-questions or at the main clause verb-Q position in main clause wh-questions. Thus there was no clear effect of local wh-scope calculation. On the other hand, there were continuing (R)AN effects in response to wh-questions. In the embedded clause region, all wh-questions elicited (R)AN effects in comparison to all yes/no-questions. Main clause wh-questions in comparison to embedded clause wh-questions elicited (R)AN effects from the embedded to the main clause verb positions. In comparison to their main clause yes/no-question counterparts, main clause wh-questions also elicited RAN effects at the embedded verb position, with a visible trend continuing through to sentence end. This pattern of results was consistent only with the wh-Q dependency hypothesis as opposed to the local scope calculation hypothesis; there thus seems to be a reliable neural processing correlate of the dependency between wh-elements and Q-particles in Japanese. Other effects found in the present experiment included lexical (P200 and N400) effects on adjectives and an N400-like effect to the main clause verb-Q complex of main clause yes/no-questions. Although the interpretation of RAN as a sentence-end wrap-up effect has been ruled out in the present experiment, the possibility that implicit prosody played a causal role in the ERP results has not. The next experiment was carried out to address this issue.

Experiment 3

Experiments 1 and 2 suggested that right-lateralized negativity indexes a dependency between Japanese wh-words and Q-particles. Other possible interpretations—such as sentence-
end wrap-up and dispreferred parse effects—can be ruled out. However, the possibility remains that the effects were due to implicit prosody. Although stimuli were presented visually in both experiments, subjects could have been reading sentences to themselves silently, and the purported silent intonation difference across conditions might have caused different ERP responses. As for prosody-related ERP effects, there has been an auditory ERP study (Steinhauer, Alter, & Friederici, 1999) that reported an N400-P600 complex reflecting a prosody-induced garden-path effect, as well as a large positive waveform at intonational phrase boundaries termed the “closure positive shift” (CPS). While this pattern of results is different from ours, it does not necessarily rule out the hypothesis that our effects were related to implicit prosody. Experiment 3 explored this possibility.

Beyond the lexical prosody level, there are said to be two levels of prosodic structure in Japanese: accentual phrase and intonational phrase (Beckman & Pierrehumbert, 1986; Beckman, 1996). The accentual phrase is defined by its tonal markings, and it is the domain of two delimitative peripheral tones, the phrasal H (high) and the boundary L (low). The intonational phrase is the domain for downstep (or catathesis), which is the compression and lowering of a pitch range following an accented phrase. Within an intonational phrase, a series of lexical accents are compressed in the form of a descending staircase. Then the pitch is reset upwards at the beginning of each new intonational phrase.

Japanese wh-words such as nani ‘what’ and dare ‘who’ are accented in wh-questions and are said to form an intonational phrase with the following word(s), resulting in their deaccentuation or pitch reduction (Maekawa, 1991a, b). Thus, when a wh-element and a verb are next to each other and the verb is lexically accented, as in Nani-ga mieru ‘What-NOM is.seen?’.
the accent on the verb disappears. On the other hand, the accent on the verb remains for yes/no-questions, as in *Nani-ka mieru* ‘Something is seen?’ (see Figure 11).

---Figure 11---

It is possible that subjects may have imagined wh-induced pitch reduction when they read wh-questions but not when they read yes/no-questions, and this may have influenced their brain responses. However, since stimuli were presented relatively slowly (i.e. every 650 ms) word by word instead of as entire sentences, it is questionable whether subjects would have created intonational phrases across individual presentations of words. One native speaker of Japanese we consulted felt that each presentation of a word became its own intonational unit. Nonetheless, there have been reports of the influence of implicit prosodic structure in ambiguity resolution with word-by-word self-paced reading times (Implicit Prosody Hypothesis: Fodor, 1998, 2002; Quinn, Abdelghany, & Fodor, 2000; also see Bader, 1998 and Hirose, 1999). In order to investigate the hypothesis that subjects created intonational phrases across presentations of words, a prosody analysis of sample sentence tokens from Experiments 1 and 2 was conducted.

**Methods**

Two native speakers of Japanese (one male and one female, Tokyo dialect) recorded 24 experimental sentences (6 sets x 4 conditions) that were pseudorandomized with 24 filler items for both mono-clausal (Experiment 1) and bi-clausal (Experiment 2) experimental stimuli. The digitized speech signals (44.1 kHz/16 bit sampling rate) were analyzed in terms of F₀ (fundamental frequency) contours. For the mono-clausal stimuli, the F₀ maxima of the verb (before the sentence final LH contour for interrogatives) for all conditions and the subject noun for the scrambled conditions were measured and compared using paired one-tailed t-tests. For the bi-clausal stimuli, the F₀ maxima of the embedded verb (the last word of the embedded
clause) and the modifier of the matrix subject (the first word of the main clause) were measured and compared using repeated measures ANOVAs (treating item as a random factor) with Tukey HSD posthoc comparisons.

Results and Discussion for Mono-clausal Stimuli

Figure 12 shows typical $F_0$ contours for the four experimental conditions for the mono-clausal stimuli. As for wh-in-situ questions (a) in comparison to yes/no-in-situ questions (b), a salient peak at ‘what/who-ACC’ and subsequent pitch reduction at the following verb position 

\[
\text{male speaker}: t(5) = 3.29, p = .011; \text{female speaker}: t(5) = 1.90, p = .065
\]

were observed. As for scrambled wh-questions (c) in comparison to scrambled yes/no-questions (d), a salient peak at ‘what/who-ACC’ and subsequent pitch reduction starting from the following word [(subject noun)]

\[
\text{male speaker}: t(5) = 4.81, p = .002; \text{female speaker}: t(5) = 5.03, p = .004
\]

to the sentence-end 

\[
\text{male speaker}: t(5) = 3.22, p = .012; \text{female speaker}: t(5) = 2.78, p = .003
\]

were observed.

Based on the above observations, the possibility that wh-in-situ-questions have a different intonational pattern from yes/no-in-situ-questions at the verb-Q position, and that this could have been indexed by increased right-lateralized negativity, cannot be denied. On the other hand, the pitch reduction with scrambled wh-questions was long-lasting: it lasted from right after the wh-words were first uttered until sentence end. If right-lateralized negativity was due to pitch reduction induced by a wh-word, we would have expected to see it start right after the wh-words in both comparisons (scrambled and in situ). However, in the scrambled comparison we saw the effect only at the two sentence-final words, rendering an intonational account of this comparison questionable at best. Moreover, if both in-situ and scrambled comparisons elicited increased
right-lateralized negativity at sentence end, it seems preferable to attribute this amplitude
difference to one factor (such as Q-dependency or scope-calculation), accounting for both
comparisons, rather than to intonation for the in-situ comparison and something else for the
scrambled comparison. Altogether, the prosody analysis of the mono-clausal stimuli revealed no
transparent, direct relationship between effects of right-lateralized negativity and intonation.

**Results and Discussion for Biclausal Stimuli**

Figure 13 shows typical $F_0$ contours of the four experimental conditions for the bi-clausal
stimuli. Recall that unlike the mono-clausal stimuli, wh-words used in the bi-clausal stimuli
were noun-modifiers instead of entire noun phrases, such as ‘what.kind.of’ and ‘what-cup-of
(how many cups of)’. Some of the non-wh-modifiers are lexically accented, while others are not.

Let us first consider a set of conditions in which a non-wh-modifier is not lexically
accented, as in the case of *san-nen-no* ‘3-year-of’ (Figure 13a, b). In this case, there was no
observable pitch reduction induced by the non-wh-modifier. On the other hand, there was a
salient peak in pitch on the wh-modifiers such as *nan-nen-no* ‘what-year-of’ (Figure 13c, d), as
well as a subsequent pitch reduction in the wh-conditions relative to the non-wh-conditions in the
embedded clause region.

When a non-wh-modifier was lexically accented, such as with *ni-hai-no* ‘2-cup-of’
(Figure 13e, f), the lexical accent on the non-wh-modifier often created a salient peak similar to
its wh-counterpart (Figures 13e, f). However, even in this case the wh-conditions often showed a
recognizable pitch reduction compared to the non-wh-conditions at the embedded verb position.

However, when ANOVAs were run on the $F_0$ maxima at the embedded verb position (the
last word of the embedded clause), the observation of pitch reduction in wh-conditions relative to
non-wh-conditions in the embedded clause showed statistical significance only for the female speaker \[male speaker: F(3, 5) = 1.12, p = .373; female speaker: F(3,5) = 11.25, p < .001,\] Tukey HSD: embedded clause wh, main clause wh < embedded clause yes/no]. At the matrix subject modifier position (the first word of the main clause), main clause wh-questions showed the lowest F₀ among the four experimental conditions for both speakers. This, however, was again statistically significant only for the female speaker \[male speaker: F(3, 5) = 1.35, p = .263; female speaker: F(1,5) = 25.70, p < .001, Tukey HSD: main clause wh < main clause yes/no, embedded clause wh, embedded clause yes/no].

In summary, consistent with the pitch reduction following a wh-word observed in the mono-clausal stimuli, there was a pattern of wh-induced pitch reduction in the scope of a wh-element in the bi-clausal stimuli as well. At the embedded verb position, wh-conditions had a lower F₀ compared to yes/no-conditions, and at the matrix subject modifier position (first word of the main clause), main clause wh-questions had the lowest F₀ among all conditions. This pattern was generally observed for both the male and the female speaker, yet was statistically significant only for the female speaker, which may have been because females tend to have wider pitch contours. It is especially interesting that main clause wh-questions showed the lowest F₀ in the main clause region, at least visually, for both speakers. This seems to suggest a correlation between wh-induced pitch reduction and wh-scope, as the main clause region is still in the interrogative scope of a wh-element for main clause wh-questions (Cf. Deguchi & Kitagawa, 2002; Hirotani, 2003).

Recall that in Experiment 2 there were continuing RAN effects in response to bi-clausal wh-questions. In the embedded clause region, all wh-questions elicited slow anterior negative potentials in comparison to all yes/no-questions. In addition, main clause wh-questions in
comparison to embedded clause wh-questions and main clause yes/no-questions elicited RAN effects from the embedded to the main clause verb positions. The domain of these RAN effects coincides with the domain of the wh-induced pitch reduction seen in our sample production data. Thus it is hard to deny altogether the possibility that the RAN effect was related to certain (tacit) prosodic properties.

However, recall that the RAN effect was already seen at the wh-modifier position in the present experiment, and that there was a salient pitch peak rather than pitch reduction at that position. Thus RAN is at least not directly related to pitch reduction itself. Also, recall that in the mono-clausal stimuli, pitch reduction in the scrambled wh-condition set in right after the scrambled wh-word, whereas the RAN effect did not appear until the final two words of the sentence. Thus there was no perfect alignment of prosody with RAN in either Experiment 1 or Experiment 2.

Moreover, how could one ever tease apart the relative contributions of syntax, semantics, and prosody in sentences like this? One could just as well claim that subjects were imagining a lower pitch in the scope of a wh-element because they had not found the Q-particle they were looking for. Then the implicit prosody effect, if it existed, could be subsumed under the effect of a subject’s searching for a Q-particle (see General Discussion).

General Discussion

Overall, the wh-Q dependency hypothesis has been well supported. Both mono-clausal and bi-clausal Japanese wh-questions elicited effects of 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. Below we will address some related issues.

**Wh-Scope Calculation**

Although we have abandoned the hypothesis that RAN indexes local wh-scope calculation at the verb-Q position, it should be noted that we have not necessarily refuted the idea that there is a processing cost involved in wh-scope calculation. The local scope calculation hypothesis states that wh-scope calculation occurs at the Q-particle position, but this need not be the case. One could argue that the scope of wh-in-situ is calculated by linking it with its corresponding Q-particle. Then scope calculation would be an on-going process rather than something that happens locally at a verb-Q position. In this way, the effects of wh-Q expectation and wh-scope calculation can be collapsed. The parser seems to be expecting and actively seeking out a Q-particle following the appearance of a Japanese wh-word, well before the Q-particle position, rather than waiting until the Q-particle position itself. This is comparable to a variety of processing models pertaining to filler-gap dependencies, including the first-resort over last-resort Strategy (Fodor, 1978; Garnsey, Tanenhaus, & Chapman, 1989), the active filler strategy (Frazier & Clifton, 1989), and the minimal chain principle (De Vincenzi, 1991), in which the parser actively tries to associate an extracted wh-element with the first possible gap position.

Related to this, Aoshima et al. (in press) likewise showed evidence for the incremental processing of Japanese wh-words. As mentioned in the Introduction, they reported a typing mismatch effect slowdown for main clause wh-questions at embedded verbs without Q-particles, even when the wh-words placed in the main clause had to be interpreted as scrambled out of the embedded clause. They also showed that Japanese readers prefer to associate a scrambled wh-
phrase with the embedded clause rather than wait for the main clause. As seen in (14), they reported a slowdown at the embedded dative NP position for the condition with a scrambled dative wh-NP in the matrix clause (14b), in comparison to the condition with no dative wh-NP (14a).

(14)

a. Control

\[
\text{どの生徒が担任に校長が図書室で司書に本を読んだと言いましたか。}
\]

\[
\text{Dono-seito-ga tannin-ni [kocho-ga toshoshitsu-de shisho-ni hon-o yonda-to] iimashita-ka}
\]

\[
\text{which-student-N  teacher-T  principal-N  library-at  librarian-at  book-A  read-that  said-Q}
\]

‘Which student said to the teacher that the principal read the book to the librarian at the library?’

b. Scrambled

\[
\text{どの生徒に担任は校長が図書室で司書に本を読んだと言いましたか。}
\]

\[
\text{Dono-seito-ni tannin-wa [kocho-ga toshoshitsu-de shisho-ni hon-o yonda-to] iimashita-ka}
\]

\[
\text{which-student-D  teacher-T  principal-N  library-at  librarian-at  book-A  read-that  said-Q}
\]

‘To which student did the teacher say that the principal read the book to the librarian at the library?’

They argued that this is the Japanese version of a filled-gap effect (Stowe, 1986), in which the parser wants to associate the filler with a possible gap position, but the position has already been filled by something else, causing a slowdown. They argued that since the projected gap-creation for the fronted wh-word takes place before any verb position, sentence structures are built incrementally prior to the critical appearance of any verbal heads.

Aoshima et al. (in press) focused on associating the scrambled dative wh-NP with the gap position, rather than with the Q-particle. They argued that the active filler strategy (Frazier & Clifton, 1989) cannot account for their data, as it would force the parser to posit a gap as soon as
ERPs to Japanese Wh-Questions

possible, namely, in the main clause before the beginning of the embedded clause, contrary to their data suggesting that the parser prefers to associate the wh-NP with the embedded clause. It should be noted, however, that the parser reportedly prefers to associate a non-wh dative NP with the main clause rather than with the embedded clause (Kamide & Mitchell, 1999, although their main clause interpretation does not call for a scrambled structure like Aoshima et al.’s). If we shift our focus from the filler-gap association to the wh-Q association, we can say that since the parser tries to associate the wh-filler with a Q-particle as soon as possible, it prefers to associate it with the embedded clause, which has the earliest possible host for a Q-particle (i.e., embedded verb). If this is the case, then something like the active filler strategy is actually the driving force behind the parser’s preference to associate the dative wh-NP with the Q-particle in the embedded clause in Aoshima et al.’s data.

Therefore, the data by Aoshima et al. (in press) and the present study all seem to point to the conclusion that the parser actively and incrementally tries to associate a Japanese wh-element with its Q-particle, just as the parser actively tries to associate an English wh-element with its gap position. This seems to suggest that the scope of Japanese wh-in-situ is calculated by the real-time, incremental formation of a long-distance dependency with its corresponding Q-particle.

**LAN vs. RAN**

In both Experiments 1 and 2, we saw effects of anterior negativity in response to wh-questions in multiple comparisons. The effects were right-lateralized in the majority of cases, although not always statistically significantly so. While they were sometimes bilateral in distribution, they were never left-lateralized. We have seen that both filler-gap dependencies and wh-Q dependencies elicit anterior negativity, often in the form of slow potentials. Slow anterior negative potentials have generally been associated with a working memory load caused by both
sentential (King & Kutas 1995; Müller, King, & Kutas 1997; Münte, Schiltz, & Kutas 1998; Fiebach et al. 2001) and non-sentential (Ruchkin, Johnson, Canoue, & Ritter 1990; Ruchkin, Johnson, Grafman, Canoue, & Ritter, 1992) linguistic stimuli. Therefore, as discussed above, it seems feasible to interpret the anterior negative slow potentials found in our study as an index of working memory load due to a dependency between a wh-element and its corresponding Q-particle in Japanese, which may well be similar to the working memory load for associating a wh-filler with its gap in wh-movement languages. However, the question remains: why RAN instead of LAN?

First, it should be kept in mind that ours is not the first nor the only study to report bilateral and/or right-lateralized slow negative potentials in response to syntactic dependencies (King & Kutas, 1995; Müller et al., 1997). In addition, although the anterior negativity effects found in the present study were often right-lateralized, they were always found bilaterally and never solely on the right side of the scalp.

However, if we are to account for the lateralization difference, it might be a good idea to review the underlying cognitive processes involved. The possibility that all Japanese dependencies elicit RAN is not very likely, as comparisons between mono-clausal Japanese questions with scrambled and in-situ wh-/non-wh- objects also elicit (L)AN (Ueno & Kluender, 2003). Thus the lateralization difference seems specifically linked to certain properties of Japanese wh-questions. In wh-movement languages, it has been argued that storage of a wh-filler in working memory and its retrieval for filler-gap assignment causes a LAN effect. More specifically, Kluender and Kutas (1993 a,b) and King and Kutas (1995) relate the effect to the processing of a filler without a grammatical function (such as “subject” and “object”) or a thematic role (such as “agent” and “patient”) until the parser hits the verb-gap position and is
able to find this missing information. Fiebach et al. (2001) claim that LAN is linked to “a syntactic working memory process that maintains the filler actively available until the gap is licensed and the syntactic chain between filler and gap can be established (p. 327)”. However, what is ambiguous about Japanese wh-elements is their interrogative scope, until they are linked with their corresponding Q-particles.

For both wh-movement and wh-in-situ languages, the parser has to hold something with missing information for some extent of time, yet the type of missing information is different. What is unresolved about a wh-filler is fairly syntactic in nature, such as grammatical function, thematic role (although this is partially semantic), and structural chain formation. On the other hand, what is unresolved about a wh-in-situ is more logico-semantic in nature, namely, its interrogative scope. Although both kinds of information are needed for the successful interpretation of wh-questions, the computation of syntactic and logico-semantic properties could involve different neural generators. This may well have contributed to the left- vs. right-lateralization differences.20

Implicit Prosody

Recall that pitch reduction on the words following wh-words was observed within the scope of the interrogative in our sample production data in Experiment 3. The domain of pitch reduction generally coincided with the sentence positions at which RAN effects were elicited. RAN was not found until later in the mono-clausal scrambled wh- vs. yes/no-question comparison in Experiment 1, and RAN was already found at the pitch peak coinciding with the wh-modifier position in Experiment 2. As mentioned in Experiment 3, it seems more plausible to use one factor (namely Q-dependency) to account for all of the RAN effects, rather than implicit prosody for some comparisons and something else for others. In addition, recall that we
have already concluded that the scope of wh-in-situ is calculated by the linkage with its corresponding Q-particle. We could thus claim that subjects were imagining a lower pitch in the scope of a wh-element because they had not found the Q-particle they were looking for. Then even if the tacit prosody effect existed, it could be subsumed under the effect of a subject’s searching for a Q-particle. If wh-scope in Japanese is calculated by associating wh-words with Q-particles (which seems perfectly plausible), and if pitch reduction reflects wh-scope, then we are just observing two different overt manifestations (brain response, pitch contours) of the same underlying cognitive phenomenon. In other words, there is no causal relationship that can be discerned here, as we could just as well argue that the brain response was driving the pitch reduction, since both (by hypothesis) indirectly reflect the establishment of wh-scope.

Summary and Conclusion

This study investigated the processing of Japanese wh-questions using ERPs. This pattern of results was consistent only with the wh-Q dependency hypothesis as opposed to the local scope calculation hypothesis. There seemed to be 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 that of Aoshima et al. (in press) suggests that the interrogative scope of a wh-word in a Japanese wh-question is licensed by the incremental formation of a long-distance dependency with its corresponding Q-particle.
Footnotes

1 Except in the case of an echo question like Calvin brought what?, designed to convey incredulity or to seek reconfirmation of correct comprehension.

2 The exception is when there is a rising intonation in speech, which can be interpreted as a prosodic version of a Q-particle.

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

4 The total number of subjects actually run was 27. However, due to experimenter error, four subjects saw stimuli at a different presentation rate from the others, and their data had to be discarded. In addition, one subject exhibited excessive muscle tension that interfered with the EEG recording, and two subjects never returned for their second session, so their data also had to be discarded.

5 The presentation rate of 650 ms per bunsetsu is longer than the standard 500 ms per word for English, but was deemed optimal after consulting three native speakers of Japanese. Similarly, Garnsey, Yamashita, Ito, and McClure (2001) used 700 ms per bunsetsu in their Japanese ERP study. Nakagome et al. (2001) used a presentation rate of 1500 ms per bunsetsu.

6 Because of this, sentence-final words did not show typical positive drifts; see Figures 2 and 3.

7 It was discovered after the data collection that dare ‘who’ is much more conventionally written with less graphically complex phonemic characters (Amano & Kondo, 2000).
The results were basically the same even when the time window was extended to 300-900 ms. Only the temporal array showed a significant main effect \( [F(1, 19) = 4.69, p = .031] \) and a marginal condition x hemisphere interaction \( [F(1, 19) = 3.40, p = .081] \) that indicated a greater right-lateralized negativity in response to wh-in-situ questions compared to yes/no-in-situ questions.

Both sentence types had five words after the scrambled objects: a subject noun phrase with three words, an adverbial, and a verb (‘that reckless adventurer-NOM finally discovered-Q’ in the sample stimuli in Table 1). However, since wh-phrases showed P200 and N350 differences relative to demonstratives, as discussed above (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 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’

Scrambling is typical for a long embedded clause in Japanese (see Yamashita, 2002).

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

A corpus count of *Asahi Shinbun* (Amano & Kondo 2000), a popular Japanese newspaper, was used to examine the token frequency of the main clause verbs used. Many of them involved a combination of a noun and a light verb suru ‘do’ (e.g., v. *tennis-suru* ‘tennis-do’ ← n. *tennis* ‘tennis’); such verb forms are typical in Japanese. Some of the verbs had the token frequency of their verb forms listed, while for others only the token frequency of their noun forms was listed. Only the pairs (‘ask’ vs. ‘say’ verbs) that had the token frequency of either
both verb forms or both noun forms listed were considered, yielding 169 pairs out of the 200 pairs. A paired t-test comparing the token frequency of the two groups of verbs [mean token frequency: 30,817 (sd 40,094) for ‘say’ verbs; 29,727 (sd 29,708) for ‘ask’ verbs] showed no significant difference between the two groups [t(168) = -.374, p = .709].

As 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)

Due to the N400 effect in response to adjectives, the 300-600 ms window was not considered to be a good latency window to test the anterior negativity to ‘what.kind.of’. The ANOVAs run in this time window showed a significant or marginal condition x anteriority interaction in all arrays [midline: F(2, 38) = 3.36, p = .08; parasagittal: F(4, 76) = 4.97, p = .013; temporal: F(2, 38) = 9.99, p = .004], indicating that non-wh-modifiers were more negative than wh-words at posterior sites (and that wh-words were more negative than non-wh-words at anterior sites for parasagittal and temporal arrays). When the mean amplitudes between 300 and 600 ms from seven frontal electrodes (Fp1/2, F7/8, F3/4, Fz) and seven posterior electrodes (O1/2, T5/6, P3/4, Pz) were submitted to an ANOVA, there was a significant condition x anteriority interaction [F(1, 19) = 7.42, p = .014]. Visually, this appeared to be due to the overlap with the N400 effect to non-wh words and the anterior negativity effect to wh-words.

On visual inspection there appeared to be a small N400-like effect to main clause wh-questions vs. main clause yes/no-questions at posterior electrodes. ANOVAs run in the window
of 350-450 ms revealed no significant or marginal differences. There also appeared to be a difference in the 0-100 ms region between the two conditions, but ANOVAs run in the window of 0-100 ms revealed no significant or marginal differences.

16 In Beckman & Pierrehumbert (1986), this level is called “intermediate phrase”.

17 Instead of an overt Q-particle, a sentence final LH contour (a prosodic version of a Q-particle) is used in this example.

18 The word immediately following, *sono* ‘the’, was integrated into the accentual phrase with the scrambled wh-/non-wh-object and was not easily measurable independently.

19 Another possibility is that at the end of the sentence, subjects repeated to themselves the entire sentence with an appropriate intonation contour, and this may have caused the sentence-final right-lateralized negativity effect for both in-situ and scrambled comparisons. However, this interpretation is questionable for two reasons. First, recall that the sentence-final effects started within 300 ms of the final verb-Q complex. It is doubtful whether subjects could recall the entire sentence that rapidly. Second, even if we assume that subjects are really fast at repeating stimulus sentences subvocally, it is doubtful whether they could read them only once and memorize them well enough to repeat them at the end of the sentence—while processing the sentences for meaning, in case there was a following comprehension question, and trying not to blink the entire time.

20 One way to further scrutinize the exact driving forces of the right-lateralization might be to investigate the processing of different types of dependencies involving wh-words and non-question-particles in Japanese, such as those with universal and existential forces (Cf. Nishigauchi, 1990).
References


Kluender, R., & Münte, T. F. (1998). *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, Newark, NJ.


Appendix A: Sample Stimuli for Experiment 1

Experimental Conditions:

a. Wh-in-situ questions

あの/幼い/男の子によると/その/奇想天外な/手品師がいきなり/何を/取出したんですか。

b. Demonstrative(yes/no)-in-situ questions

あの/幼い/男の子によると/その/奇想天外な/手品師がいきなり/それを/取出したんですか。

c. Scrambled wh-questions

あの/幼い/男の子によると/何を/その/奇想天外な/手品師がいきなり/取出したんですか。

d. Scrambled demonstrative (yes/no) questions

あの/幼い/男の子によると/それを/その/奇想天外な/手品師がいきなり/取出したんですか。

Below are the a-versions only.

1a. あの/幼い/男の子によると/その/奇想天外な/手品師がいきなり/何を/取出したんですか。

2a. あの/デパートの/店員によると/その/げたくなる/社長婦人が/最初に/何を/選んだんですか。

3a. あの/上品な/奥方によると/その/後任の/女中さんが/てきぱき/何を/乾かしたんですか。

4a. あの/とんかつ屋の/経営者に/よると/その/だらしない/アルバイト学生が/とうとう/何を/返したんですか。

5a. あの/けんもほろろの/役人によると/その/社会派の/ジャーナリストが/堂々と/何を/書いたんですか。

6a. あの/担任の/教員によると/その/悪童の/小学生が/こそそこ/何を/隠したんですか。

7a. あの/ブティックの/ハウススマヌカンによると/その/大スターの/タレントが/こっそり/何を/買ってたんですか。

8a. あの/ごっついつい/長男によると/その/可愛らしい/妹が/一生懸命/何を/飾ったんですか。

9a. あの/販売部の/営業主任に/よると/その/不注意な/派遣社員が/またまた/何を/見落としたんですか。

10a. あの/大柄な/現場監督によると/その/そっかしい/作業員が/うっかり/何を/切ったんですか。

11a. あの/近くの/いたずらっ子達によると/その/ブチの/小犬が/偶然/何を/踏んだんですか。

12a. あの/農家の/おかみさんに/よると/その/乱暴な/少年達が/どうして何を/踏んだんですか。

13a. あの/秘書室の/みんなによると/その/新参の/秘書が/どうにか/何を/まとめたんですか。

14a. あの/町内会の/副班長によると/その/おとなしい/選手が/急に/何を/見たんですよ。
15a. あのローカルな新聞によると、その命知らずの冒険家がついに何を発見したんですか。
16a. あの同室の入院患者によると、その長患いの病人が、あらかじめ何を飲んだんですか。
17a. あの当直の医師によると、その軽率な糖尿病患者が、知らぬ間に何を食べたんですか。
18a. あの新興宗教の教団によると、そのいんちきくさい霊媒が、どんな何を使ったんですか。
19a. あのお固い美術委員会によると、その意欲的な芸術家が、一月で何を、作ってたんですか。
20a. あのライバル支店の課長代理によると、そのすご腕のセールスマンが、すんなり何を、売ったんですか。
Appendix B: Sample Stimuli for Experiment 2

Experimental Conditions:

a. Embedded clause wh-questions

専務が/何台の/パソコンを/買ったか/経理課の/係長が/尋ねましたか。

b. Matrix clause wh-questions

専務が/何台の/パソコンを/買ったと/経理課の/係長が/言いましたか。

c. Embedded clause yes/no-questions

専務が/二台の/パソコンを/買ったか/経理課の/係長が/尋ねましたか。

d. Matrix clause yes/no-questions

専務が/二台の/パソコンを/買ったと/経理課の/係長が/言いましたか。

Below are a- and d-versions only.

1a. 専務が/何台の/パソコンを/買ったか/経理課の/係長が/尋ねましたか。

1d. 専務が/二台の/パソコンを/買ったと/経理課の/係長が/言いましたか。

2a. 犯人が/どんな/バッグを/盗んだか/担当の/刑事が/質問しましたか。

2d. 犯人が/ブランド物の/バッグを/盗んだと/担当の/刑事が/報告しましたか。

3a. 園児達が/どんな/野菜を/食べ残したか/保育園の/保母さんが/聞きましたか。

3d. 園児達が/かたい/野菜を/食べ残したと/保育園の/保母さんが/言いましたか。

4a. お客さんが/何杯の/コーヒーを/頼んだか/喫茶店の/店長が/尋ねましたか。

4d. お客さんが/二杯の/コーヒーを/頼んだと/喫茶店の/店長が/話しましたか。

5a. 会長が/どんな/歌を/歌ったか/現場の/職員が/問い掛けましたか。

5d. 会長が/流行の/歌を/歌ったと/現場の/職員が/ささやきましたか。

6a. 教授が/どんな/本を/出版したら/研究室の/助手が/思い出しましたか。

6d. 教授が/雑学の/本を/出版したら/研究室の/助手が/つぶやきましたか。

7a. サッカー選手が/何年の/契約を/交わしたか/番組の/レポーターが/質問しましたか。

7d. サッカー選手が/三年の/契約を/交わしたと/番組の/レポーターが/報告しましたか。

8a. 部長が/どんな/機械を/壊したか/現場の/工事長が/確認しましたか。

8d. 部長が/大切な/機械を/壊したと/現場の/工事長が/断言しましたか。

9a. 評論家が/何人の/作家を/誉めたか/文芸雑誌の/編集長が/尋ねましたか。

9d. 評論家が/二人の/作家を/誉めたと/文芸雑誌の/編集長が/語りましたか。

10a. お手伝いさんが/どんな/お皿を/割ったか/お屋敷の/奥さんが/聞きましたか。
10d. お手伝いさんが大切なお皿を割ったとお屋敷の奥さんが決付けましたか。

11a. 老人が万円の現金をなくしたか派出所の警察官が調査しましたか。
11d. 老人が万円の現金をなくしたと派出所の警察官が書き留めましたか。

12a. 魚屋さんがどんな魚を仕入れたか町内の寿司屋さんが聞出しましたか。
12d. 魚屋さんが新鮮な魚を仕入れたと町内の寿司屋さんが口外しましたか。

13a. 小学生が何冊の本を読んだかアルバイトの塾講師が尋ねましたか。
13d. 小学生が二冊の本を読んだとアルバイトの塾講師が伝えましたか。

14a. 新入社員がどんな書類をなくしたか総務部の事務員が問詰めましたか。
14d. 新入社員が重要な書類をなくしたと総務部の事務員が口外しましたか。

15a. 作業員が数箱を運んだか当直の係員が確認しましたか。
15d. 作業員が四つの箱を運んだと当直の係員が証言しましたか。

16a. 息子が何人の女性を愛したか田舎の母親が気付きましたか。
16d. 息子が二人の女性を愛したと田舎の母親が打明けましたか。

17a. 秘書がどんな資料を捨てたか役員室の室長が調査しましたか。
17d. 秘書が極秘の資料を捨てたと役員室の室長が主張しましたか。

18a. おばあさんが何本のバラを植えたか近所の係が聞きましたか。
18d. おばあさんが三本のバラを植えたと近所の係が言いましたか。

19a. 一年生がどんな花瓶を割ったか中学校の教頭先生が見破りましたか。
19d. 一年生が高価な花瓶を割ったと中学校の教頭先生が決め付けましたか。

20a. 生徒が一つの選択科目を勉強したか臨時雇いの家庭教師が確認しましたか。
20d. 生徒が三つの選択科目を勉強したと臨時雇いの家庭教師が報告しましたか。
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

Types in modifiers to the embedded object noun phrase

<table>
<thead>
<tr>
<th></th>
<th>Wh-modifiers</th>
<th>Non-wh-modifiers</th>
<th>Number of stimulus sets</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adjective pairs</td>
<td>donnna 'what.kind.of'</td>
<td>adjectives (e.g. 'new')</td>
<td>100</td>
</tr>
<tr>
<td>Numeral classifier pairs</td>
<td>wh-classifiers (e.g., 'what-cup-of (how many cups of)')</td>
<td>non-wh-classifiers (e.g., '2-cup-of')</td>
<td>100</td>
</tr>
</tbody>
</table>
Table 3

Sample stimuli for Experiment 3

a. Embedded clause wh-questions

[senmu-ga donna pasokon-o katta-ka] keiri-no kakaricho-ga__kikimashita-ka.
director-N what.kind.of PC-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

[senmu-ga donna pasokon-o katta-to] keiri-no kakaricho-ga__imashita-ka.
director-N what.kind.of PC-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

[senmu-ga atarashii pasokon-o katta-ka] keiri-no kakaricho-ga__kikimashita-ka.
director-N new PC-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

[senmu-ga atarashii pasokon-o katta-to] keiri-no kakaricho-ga__imashita-ka.
director-N new PC-A bought-that accounting-of manager-N said.POL-Q

‘Did the accounting manager say the director bought a new computer?’
Figure Captions

Figure 1. ERPs (n = 20) at a left frontal electrode (F3) 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’) of wh-in-situ vs. yes/no-in-situ-questions.

Figure 3. ERPs (n = 20) at midline and parasagittal electrodes (Fp1/2, F3/4, C3/4, P3/4, O1/2, Fz, Cz, Pz) relative to a 400 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’) of scrambled wh- vs. yes/no-questions.

Figure 4. ERPs (n=20) recorded from all electrodes relative to a 100 ms prestimulus baseline. Shown are the ERPs to ‘what.kind.of’ vs. adjectival modifiers.

Figure 5. ERPs (n=20) recorded from all electrodes relative to a 100 ms prestimulus baseline. Shown are the ERPs to wh-numeral ‘what.kind.of’ vs. non-wh-numeral ‘2-cup-of’ classifiers.

Figure 6. ERPs (n=20) recorded from all electrodes relative to a 200 ms prestimulus baseline. Shown are the ERPs to the embedded object position (‘what.kind.of/new PC-ACC’ in Table 3) of all wh- vs. yes/no-questions.

Figure 7. 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 embedded vs. main clause wh-questions.

Figure 8. ERPs (n=20) recorded from all electrodes relative to a 100 ms prestimulus baseline. Shown are the ERPs to the embedded verb position (‘bought-that’ in Table 3) of main clause wh-vs. yes/no-questions.
Figure 9. ERPs (n=20) recorded from all electrodes relative to a 400 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 Table 3) of embedded vs. main clause wh-questions.

Figure 10. ERPs (n=20) recorded from all electrodes relative to a 400 ms prestimulus baseline. Shown are the ERPs to the region from the embedded to the main clause verb position (‘bought-that accounting-of manager-NOM asked-Q’ in Table 3) of main wh- vs. yes/no-questions.

Figure 11. The F0 (fundamental frequency) contours of the wh-question Nan-ga mieru ‘What-NOM is.seen?’ (left) and the yes/no-question Nanika mieru ‘Something is.seen?’ (right), as uttered by a male Tokyo Japanese speaker. (adopted from Maekawa, 1991a: 203)

Figure 12. The F0 contours of the last six words of (a) a wh-in-situ question, (b) a yes/no-in-situ question, (c) a scrambled wh-question, and (d) a scrambled yes/no-question, as uttered by a male Tokyo Japanese speaker.

Figure 13. The F0 (fundamental frequency) contours of embedded clause wh-questions (c and g), main clause wh-questions (d and h), embedded clause yes/no-questions with unaccented (a) and accented (e) modifiers, and main clause yes/no-questions with unaccented (b) and accented (f) modifiers, as uttered by a female Tokyo Japanese speaker.
Figure 1

![Graph showing ERP responses to Japanese Wh-questions.]

- **N350**
- **P200**

---

...... who-ACC  誰を
...... what-ACC  何を
--- that-person-ACC  あの людей를
--- that-ACC  それを
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。

Wh-in-Situ:  ..that reckless adventurer-N finally what-A discovered-Q.

Yes/No-in-Situ:  ..that reckless adventurer-N finally that-A discovered-Q.
Figure 3

**Scrambled Wh:**
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。

**Scrambled Yes/No:**
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。

**Scrambled Wh:**
...what-A that reckless adventurer-N finally discovered-Q.

**Scrambled Yes/No:**
...that-A that reckless adventurer-N finally discovered-Q.
Figure 4

what.kind.of どんな

new 新しい
Figure 5

............. what-cup-of (how many cups of)  何杯の

______ 2-cup-of  二杯の
Figure 6

小事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-N what.kind.of PC-A bought-Q/that] accounting-of manager-N __ asked/said-Q.

——— All Wh

小事務/幹事 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-N new PC-A bought-Q/that] accounting-of manager-N __ asked/said-Q.
Figure 7

........... Embedded Wh
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。
[Director-N what.kind.of PC-A bought-Q] accounting-of manager-N __ asked-Q.

........... Main Wh
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。
[Director-N what.kind.of PC-A bought-that] accounting-of manager-N __ said-Q.
Figure 8

Main Wh
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。
[Director-N what.kind.of PC-A bought-that] accounting-of manager-N said-Q.

Main Yes/No
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。
[Director-N new PC-A bought-that] accounting-of manager-N said-Q.
Figure 9

………… Embedded Wh
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。
[Director-N what.kind.of PC-A bought-Q] accounting-of manager-N __ asked-Q.

………… Main Wh
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。
[Director-N what.kind.of PC-A bought-that] accounting-of manager-N __ said-Q.
Figure 10

Main Wh
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。
[Director-N what.kind.of PC-A bought-that] accounting-of manager-N __ said-Q.

Main Yes/No
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。
[Director-N new PC-A bought-that] accounting-of manager-N __ said-Q.
Figure 11
Figure 12

(a) Wh-in-situ-question

(b) Yes/no-in-situ-question

(c) Scrambled wh-question

(d) Scrambled yes/no-question
Figure 13

(a) Embedded yes/no (w/ unaccented modifier)    (b) Main yes/no (w/ unaccented modifier)

(c) Embedded wh-question    (d) Main wh-question

(e) Embedded yes/no (w/ accented modifier)    (f) Main yes/no (w/ accented modifier)

(g) Embedded wh-question    (h) Main wh-question