MEG responses in the comprehension of Japanese sentences*

Hirohisa Kiguchi and Edson T. Miyamoto
KANAZAWA INSTITUTE OF TECHNOLOGY AND U. TSUKUBA

1. Introduction

Two types of event related potentials (ERPs) have often been used in the electroencephalography (EEG) literature in order to investigate brain activity during sentence comprehension. One is the P600, which peaks around 600 ms after the beginning of a word that is syntactically unexpected (Osterhout & Holcomb 1992). The second is the so-called N400, peaking around 400 ms after the onset of a word that is semantically anomalous in the sentence context (Kutas & Hillard 1980). We report the preliminary analyses of a study investigating the corresponding responses in magnetoencephalography (MEG) using Japanese sentences.

There are two primary goals. First, we would like to replicate previous MEG findings concerning the localization of the N400 (Helenius et al. 1998, Halgren et al. 2002) and address some methodological issues in those studies. Although they do not seem to pose a serious problem for the N400, these methodological issues may nevertheless be critical for later brain responses such as the P600, for which localization has been claimed to be more problematic (Friederici 2003). By adopting more careful methodologies, we increase our chances of success in our second goal, namely the investigation of the P600 in MEG. For clarity sake, we divide this paper along those two goals by reporting our work in two parts. In the first part, we summarize previous literature and then
report our results on the N400; in the second part, we provide the background and the results concerning the P600.

2. **The first part of our experiment**

Semantically inappropriate words elicit a negative EEG component peaking at about 400 ms after stimulus onset (the N400 response; Kutas & Hillyard 1980, 1983). For example, the sentence in (1a) exhibits an N400 response at the semantically inappropriate word *mustard*, relative to *radio* in (1b). (Semantically anomalous examples will be preceded by a ‘#’ mark.)

(1)  

a. # He got in the car and turned on the mustard.  
    b. He got in the car and turned on the radio.

Although EEG has been useful to characterize the sensitivity of the N400 to the semantic properties of the stimuli, the scalp distribution thus obtained has not been as helpful in the localization of the generators of this component because electrical potentials are warped when traveling through brain tissue and bone, thus complicating the localization of responses (e.g. Hämäläinen 1995, Phillips et al. 1995).

Magnetoencephalography (MEG) detects the magnetic field generated by cortical transsynaptic current flows. Since magnetic signals, unlike electrical potentials, are not distorted by brain tissue or bone, localization of the currents underlying the activity is more realistic in MEG than in EEG.

In previous studies using MEG, N400 related effects have been reported to be localized in or near the left superior temporal cortex. Helenius et al. (1998) conducted a study in Finnish comparing the sentence-final words in semantically anomalous sentences such as ‘The pizza was too hot to sing’ to those in semantically appropriate sentences such as ‘The piano was out of tune’ in order to
localize equivalent current dipoles (ECD) for the N400 generator. Eight out of ten participants showed distinct activation in the left superior temporal cortex. Five participants showed an additional source around the posterior end of the sylvian fissure.

In another MEG study, all ten subjects showed distinct activation in or near the left superior temporal cortex in response to semantically inappropriate words in sentence-final position of English sentences compared to predictable semantically appropriate words (Halgren et al. 2002).

2.1 Some methodological issues

There are two basic methodological issues that need to be considered in those earlier studies. First, the critical word in the semantically anomalous sentences differed from the control word in the semantically acceptable sentences. In fact, in some cases the entire sentences compared were different (e.g., Helenius et al. 1998). Ideally, we would like to keep the sentences compared as similar as possible. Unfortunately, it is not advisable to show exactly the same sentence in its appropriate (e.g., the piano was out of tune) and its semantically anomalous version (e.g., the truck was out of tune) to the same participant out of concerns that repeated exposures may change the nature of the response. Thus, in the present experiment, the critical word was kept constant across conditions as comparisons were made for the same verb (see next section). In order to avoid repeated exposures to the same sentence context, we adopted a Latin Square design in which each participant sees each sentence context once, and presentation is counterbalanced across participants so that each context seen in each of the experimental conditions the same number of times overall. Latin Square does not resolve the issue for within-subject comparisons, but we believe it is an improvement compared to earlier presentation modes.

Second, the critical comparison in previous MEG studies was conducted for
words placed in sentence-final position. It has been pointed out that comparisons at sentence end can affect ERP waveforms because of end-of-sentence wrap-up effects and preparatory processes for the response to the comprehension task that usually follows each sentence. (Hagoort, Brown & Osterhout 1999, de Vincenzi et al. 2003). For example, sentence-ending words are followed by a large positive wave form not typically observed following words at the end of embedded sentences (Friedman et al. 1975). Furthermore, final words in unacceptable sentences elicit an enhanced N400-like effect, regardless of whether the unacceptability is semantic or syntactic (e.g., Hagoort et al. 1993, Osterhout & Holcomb 1992, 1993). Therefore, it is conceivable that MEG recordings taken at sentence end could at least partially be obscured by end-of-sentence processes. As discussed in the next section, the critical word in our stimuli is the verb in embedded clauses, which is followed by words in the matrix clause, therefore it avoids sentence-end problems.

2.2 Material and Methods

Stimuli: Japanese is a head-final language; in particular, the verb is always the last constituent of the clause as can be seen in the following example.

(2)  # Masao-ga hon-o nonda.
     Masao-Nom book-Acc drank
     ‘Masao drank a book.’

In order to avoid target words in sentence final position, the clause containing the semantically mismatched verb and object can be embedded within another clause as follows.
However, this type of sentence is known to cause considerable confusion because it involves a self-embedding configuration and contains two consecutive nominative-marked NPs (see Miyamoto 2002, and references therein). To avoid those problems, the embedded clause was displaced to the beginning of the sentence as indicated below (the marker $t_1$ indicates the original position of the embedded clause shown between square brackets; see Tsujimura 1999, for a general introduction to Japanese syntax in generative grammar).

(4) # [Masao-ga hon-o nonda-to]$_1$ John-ga $t_1$ omotta.
    ‘John thought that Masao drank a book.’

An example pair of sentences used in the experiment can be seen in (5). The sentence in (5a) is semantically anomalous, and (5b) is its corresponding control.

(5) a. # Masao-ga hon-o kissaten-de nonda-to Jiro-ga omotta-no?
    Masao-Nom book-Acc café-at drank-Comp Jiro-Nom thought-QP
    ‘Did Jiro think that Masao drank a book at the café?’

b. Masao-ga mizu-o kissaten-de nonda-to Jiro-ga omotta-no?
    Masao-Nom water-Acc café-at drank-Comp Jiro-Nom thought-QP
    ‘Did Jiro think that Masao drank water at the café?’

The blank spaces indicate the segmentation used in the experimental presentation. The critical region (or bunsetsu, that is, a content word followed by
functional material) to be used for comparison is the embedded verb nonda ‘drank’ together with the complementizer to ‘that’ (the complementizer is more naturally shown together with the verb, therefore all comparisons were conducted considering them as a single unit). The only difference between (5a,b) is in the accusative-marked direct object (‘book’ or ‘water’), which is followed by the same phrase (‘at the café’) so that the critical region is not contaminated by differences in the lexical material in the earlier region. Care was taken so that both types of sentences are semantically coherent before the embedded verb. The semantic inappropriateness in (5a) is only detected at the embedded verb, which clashes with the direct object read earlier.

This type of design has a number of advantages. First, the exact same word is compared across conditions, namely the embedded verb ‘drank’. Second, possible sentence-end contaminations are avoided by placing the critical word in mid-sentence. Third, the comparison is conducted at the verb, whereas earlier work for the most part has employed nouns as critical words (but see de Vincenzi et. al. 2003, for comparisons at the verb in an EEG experiment in Italian). Finally, because of the SOV word order in Japanese (that is, Subject-Object-Verb), as opposed to the SVO order usually encountered in English and other West-European languages, there is an interesting question that we can address, namely whether the expectation created by a sequence of noun phrases (NPs; that is, the subject and the object in (5)) is qualitatively similar to the expectation that generates the N400 observed in SVO languages, in which it is the verb for the most part that determines, through selectional restriction, the types of constituents that are likely to follow. It is known from previous EEG literature in Japanese that a subject-object sequence is able to create an expectation strong enough to generate a N400 at an inappropriate ensuing verb (Nakagome et al. 2001). By localizing the N400 response in Japanese, we can further determine whether there
are qualitative differences depending on the type and order of the constituents involved.

Each participant saw a total of 50 semantically inappropriate and 50 semantically appropriate sentences presented visually, interspersed randomly with 200 similar bi-clausal sentences (see the second part of the experiment for further details).

**Subjects and Procedure:** Five right-handed Japanese native speakers (1 female, 4 males: 22-24 years old) participated after giving their informed consent. The sentences were presented using PsyScope (Cohen et al.1993) in randomized order. There was a total of six types of sentences and each participant read 300 sentences (50 tokens of each type) presented visually one bunsetsu (a content word plus functional particles) at a time in the center of the screen. The spaces in (5) indicate the segmentation adopted. Each busetsu was presented for 700 ms with ISI of 50 ms using Japanese characters with the monospaced Osaka Tohaba font. A comprehension task followed on a new screen after each sentence. Results will be reported for all sentences, regardless of comprehension response.

**Magnetic measurement:** The magnetic activity was measured in a magnetically shielded room with a 160-channel whole-head magnetometer (Kanazawa Institute of Technology, Japan). Data were sampled at 500Hz, with acquisition between 0.03 and 100Hz. Responses to target words were averaged by stimulus condition after rejecting trials with eyeblinks or other artifacts. Following averaging, data were lowpass-filtered with a cutoff frequency of 60Hz.

### 2.3 Results for the first part of the experiment

Here we report results which replicate previously reported N400-like responses in MEG. Figure 1 illustrates the overall MEG activity for 159 channels, which resulted from removing one noisy channel, of subject 1 at the embedded verb and complementizer for the semantically anomalous condition and its control.
In the semantically anomalous condition, the (ongoing) peak is observed at 372ms from the onset of the semantically anomalous embedded verb, relative to the control. The magnetic distribution of this peak is illustrated in Figure 2.

Fig.1. The averaged MEG waveforms for subject 1 at the embedded verb and complementizer for the two conditions: the semantically anomalous (top) and its control (bottom).
The root mean square (RMS) of the activity for the semantically inappropriate words (M = 49.9965 fT) including only channels used to calculate the dipoles (see the next section) was reliably larger than the RMS for the semantically appropriate words (M = 40.6270 fT) when averaged over the time window between 370-390 ms for subjects 1 to 4, and 420-440 ms for subject 5 (F(1,4)=9.94, p<0.05).

2.4 Source analysis for the N400 generator

Equivalent current dipoles (ECD) around the peak evoked by semantically anomalous words were calculated and superimposed on each subjects’ magnetic resonance image (MRI). The ECD around the peak in each subject was located in or near the left superior temporal cortex (see Figures 3 to 7).

For subject 1, dipoles from 370 ms to 380 ms including all channels were calculated, yielding goodness-of-fit (GOF) of 57.22% to 74.96%. For the remaining subjects, 43 to 55 contiguous channels from the left hemisphere were used for dipole modeling. The GOF was between 81.67% and 95.40%. For subjects 2 to 4, the N400 time window for dipole modeling was set to 380ms (370ms-390ms). For subject 5, the N400 time window was 430ms (426ms-436ms).
Fig. 3. The squares indicate the ECD location at 372 ms from the onset of the semantically anomalous embedded verb for subject 1.

Fig. 4. The squares indicate the ECD location at 370 ms from the onset of the semantically anomalous embedded verb for subject 2.

Fig. 5. The squares indicate the ECD location at 380 ms from the onset of the semantically anomalous embedded verb for subject 3.
Fig. 6. The squares indicate the ECD location at 384 ms from the onset of the semantically anomalous embedded verb for subject 4.

Fig. 7. The squares indicate the ECD location at 432 ms from the onset of the semantically anomalous embedded verb for subject 5.

2.5 Discussion

We replicated the MEG brain response to semantically anomalous words obtained in previous studies (Halgren et al. 2002, Helenius et al. 1998). For four subjects, an N400-like effect was detected peaking around 370 ms after the onset of the embedded verb in (5a), which is not compatible with the initial segment of the sentence, in comparison to the control (5b). For the fifth subject, a similar response was found around 420 ms. The results replicate the localization of the
N400 around the left superior temporal region, that has been reported in previous studies (Halgren et al. 2002; Helenius et al. 1998).

The similarity in the localization of the response suggests that a sequence of NPs leading to a mismatched verb generates the same type of response as a verb leading to a mismatched NP. This type of result, if confirmed by future more detailed investigations, may support earlier proposals in generative grammar according to which *heads* (verbs, in particular) do not hold a special status in *subcategorization*, in the sense that this is a symmetrical relation in which heads subcategorize as well as are subcategorized by the context (e.g., the surrounding NPs) in which they occur (Chomsky, 1965). Although not incompatible, the present result would not be as straightforwardly explained within endocentric approaches that have dominated much of generative linguistics since the 1980’s (see for example Haegeman 1991, Sag and Wasow 1999, for overviews).

The fact that we replicated previous N400 localizations validates the methodological changes that we have adopted (same critical word across conditions in non-final position with sentences distributed in a Latin Square design). The advantage is that the present methods are less likely to yield contaminated waveforms, which is particularly important for the second part of the paper where we investigate the P600.

3. **The second part of the experiment**

In the second part of our experiment, we investigate P600 effects with wh-phrases in Japanese using MEG.

3.1 **Wh-phrases in Japanese**

The canonical position of wh-phrases (e.g., *what, who, how often, what type of computer*) in Japanese is in-situ, that is, unlike English, Japanese wh-phrases need not be fronted to the beginning of the clause and can be pronounced in the
same position as regular (that is, non-wh) noun phrases (NP). This could suggest that Japanese wh-phrases and regular NPs are processed in exactly the same manner. However, Japanese and other wh-in-situ languages have been claimed to require wh-phrases to be accompanied by some morphological marking (Cheng 1991; see Nishigauchi 1990 for a detailed discussion on Japanese).

Consider the examples in (6). As the translations in English indicate, the scope of the wh-phrase in (6a) is the embedded clause, whereas the scope of the wh-phrase in (6b) is the matrix clause.

(6) a. Senmu-ga donna pasokon-o tsukatteiru-ka kakaricho-ga
director-Nom what-kind-computer-Acc using-is-QP supervisor-Nom
kiita-no?
asked-QP
‘Did the supervisor ask what kind of computer the director is using?’

b. Senmu-ga donna pasokon-o tsukatteiru-to kakaricho-ga
director-Nom what-kind-computer-Acc using-is-that supervisor-Nom
iitta-no?
said-QP
‘What kind of computer did the supervisor say that the director is using?’

As can be seen in the English translations, the position where the wh-phrase is pronounced in this language determines its scope. In contrast, the Japanese wh-phrase *donna pasokon* ‘what-kind-computer’ in (6a,b) is pronounced in the exactly same position inside the embedded clause. It is the position of the question particle (QP) that determines the scope of the wh-phrases in Japanese. More specifically, the scope of the wh-phrase in (6a) is determined by the position of the QP *ka* at the end of the embedded clause (although judgements vary, (6a) is
sometimes considered to be ambiguous as the matrix scope interpretation is also accepted, although strongly dispreferred), whereas in (6b), the QP no dictates that the scope of the wh-phrase must be the matrix clause.

3.2 Typing mismatch effects in Japanese

Given that wh-phrases in Japanese require a QP for scope purposes, Miyamoto and Takahashi (2002) proposed that wh-phrases in this language should create the expectation for a QP that they can be associated with (in the same way that wh-phrases in English-like languages create the expectation for a gap, that is an empty position in the sentence, in order to receive a semantic role; Crain & Fodor 1985, Stowe 1986). Furthermore, assuming that the required QP is posited at the earliest possible position because of working memory constraints, there should be a typing mismatch effect taking place at the embedded complementizer position of sentences like (6b), because that is the earliest position that could contain the QP, but contrary to readers’ expectations a declarative complementizer fills that position. In other words, after detecting the wh-phrase, native speakers of Japanese expect the clause to be typed as interrogative (as indicated by a QP), but instead the embedded clause in (6a) turns out to be declarative (as indicated by the declarative complementizer to), hence there is a mismatch between the expected and the actual clausal types.

Supporting those assumptions, readers were found to slow down when they read tsukatteiru-to ‘using-is-that’ in sentences like (7) (Miyamoto & Takahashi, 2002).
To sum up, according to Miyamoto and Takahashi (2002) wh-phrases in Japanese, as is the case for wh-phrases in English, generate the expectation for a specific type of constituent at the earliest possible position in the sentence and slow reading times are observed if that expectation is violated.

3.3 An MEG investigation of typing mismatch effects

The goal of the second part of the experiment is to investigate magnetic activities in the brain for typing mismatch effects, which were behaviorally observed by Miyamoto and Takahashi (2002). We predict that typing mismatch effects elicit a P600 response, based on Kaan et al.’s (2000) claim that the P600 is not just a marker of ungrammaticality but rather a reflection of syntactic integration difficulty.

Note that the sentence in Japanese in (7) is not ungrammatical as the requirement for a QP in order to determine the scope of the wh-phrase is satisfied at sentence end. It is only the case that the declarative complementizer to following the embedded verb violates readers’ expectations for a QP at that point. The unexpectedness of the declarative complementizer should cause difficulty in attaching it to the grammatical representation of the sentence fragment read up to that point. Hence, we predict that a typing mismatch effect should elicit a P600 response.
*Stimuli:* A total of 50 target sentences with matrix-scope whs as in (8a), 50 embedded-scope wh sentences as in (8b), which satisfy the expectation for the QP at the embedded clause level, and 50 non-wh sentences as in (8c), which do not have a wh-phrase, were visually presented interspersed with 150 similar bi-clausal filler sentences.

(8) a. Masao-ga nani-o kissaten-de nonda-to Jiro-ga omotta-no?
Masao-Nom what-Acc café-at drank-Comp Jiro-Nom thought-QP
‘What did Jiro think that Masao drank at the café?’
b. Masao-ga nani-o kissaten-de nonda-ka Jiro-ga tazuneta-no.
Masao-Nom what-Acc café-at drank-QP Jiro-Nom asked-QP
‘Did Jiro asked what Masao drank at the café?’
c. Masao-ga mizu-o kissaten-de nonda-to Jiro-ga omotta-no?
Masao-Nom water-Acc café-at drank-Comp Jiro-Nom thought-QP
‘Did Jiro think that Masao drank water at the café?’

3.5 Preliminary Results

Larger activity (as measured by the root mean square, RMS, over 49 channels from the left temporal area shown in Figure 8; see also Figure 9) was detected between 600 ms and 632 ms after the onset of the embedded verb in (8a) (M=54.2 fT) in comparison to (8b) (M=29 fT; F1(1,4)=10.39, P<0.05) and (8c) (M=38.9 fT; F1(1,4)=7.26, P<0.06).
Fig. 8: Sensors of interest: 49 channels from the left temporal region.

Fig. 9. RMS from the sensors of interest for each condition averaged over five subjects’ data. Larger activity was detected peaking around 620 ms after the onset of the embedded verb in (8a) (curve with circles) in comparison to (8c) (curve with triangles) and (8b) (curve with squares).
3.6 Discussion

In the second part of our experiment, we detected larger activity approximately 600 ms after the onset of the verb-complementizer region in (8a), providing supporting evidence for the claim that wh-phrases in Japanese, as is the case for wh-phrases in English, generate the expectation for a specific type of constituent at the earliest possible point in the sentence and that difficulty ensues if that expectation is violated (Miyamoto & Takahashi 2002). The result corroborates behavioral studies indicating that readers have difficulty at the verb-complementizer region in (8a) because the wh-phrase *donna* creates the expectation for a QP (e.g., *ka*) and this expectation is contradicted by the declarative complementizer *to* ‘that’ leading to a typing mismatch effect. The result is also compatible with earlier EEG findings according to which the P600 is not just an ungrammaticality marker, but rather an indicator of difficulty in attaching an incoming constituent to the grammatical representation of the sentence fragment read so far (Kaan et al. 2000).

One of the remaining tasks in our ongoing project is the localization of the P600-like response that we observed in this experiment. Thus far, we have not been able to calculate the ECDs because of the limited distribution of the channels for which we found the larger magnetic activation to the typing mismatch effect in (8a) compared to the controls in (8b, c), and because of individual variance, for example, one subject’s activation for (8a) at the relevant channels was weaker than those for the controls. A new analysis is being conducted including only the items for which participants gave correct responses in the comprehension task.

Nevertheless, the RMS results reported above indicate that the effects are likely to lie in the left temporal region for the typing mismatch paradigm. Similarly, an fMRI study by Friederici et al. (2003) observed greater activation in the posterior superior temporal gyrus for both semantic and syntactic anomalous conditions, presumably as a result of the increased effort involved in integrating an
anomalous word into the structure being built for a sentence, although these authors reported bilateral activation of the relevant brain area.

4. Concluding remarks

We reported the preliminary analyses of a MEG study using Japanese sentences. First we investigated N400-like responses free from possible sentence-end contaminations. The results replicate the brain response to semantically anomalous words reported for previous MEG experiments conducted by Helenius et al. (1998) and Halgren et al. (2002), according to which a generator of the N400 response is located around the left temporal region. Second, we detected a P600-like response in the left temporal region in sentences in which the expectation for a question particle evoked by the presence of a wh-word is not satisfied as early as possible, providing supporting evidence for the claim that wh-phrases in Japanese generate the expectation for a specific type of constituent at the earliest possible point in the sentence (Miyamoto & Takahashi, 2002). The result is compatible with earlier EEG findings according to which the P600 is not just an ungrammaticality marker, but rather an indicator of difficulty in attaching an incoming constituent to the grammatical representation of the sentence fragment read so far (Kaan et al. 2000).

Notes

* We would like to thank Yoshiaki Adachi, Nobuhiko Asakura, Natsuko Hatsusaka, Masanori Higuchi, Jun Kawai, Shinji Kawashima, David Poeppel, Takafumi Sasaoka, Yoshihide Tamori, Jeff Walker, Tatsuya Yoshizawa, Kazuo Yamaguchi, Naoya Yokoyama. This work was partially supported by the Gakujuyutsu frontier awarded to the Human Information System Laboratories at the Kanazawa Institute of Technology.
References


Halgren, E., P. D. Rupali, N. Christensen, C. Van Patten, K. Marinkovic, J.


(kiguchih@his.kanazawa-it.ac.jp)

(miyamoto@alum.mit.edu)