An ERP study of sensory mismatch expressions in Japanese

Tsutomu Sakamoto, a, * Kana Matsuishi, b Hiroshi Arao, c and Junri Oda b

a Department of Linguistics, Faculty of Humanities, Kyushu University, Fukuoka, Japan
b Department of Psychology, Faculty of Letters, Kyushu University, Fukuoka, Japan
c Department of Psychology, Faculty of Humanities, Niigata University, Niigata, Japan

Accepted 23 July 2002

Abstract

Combinations of different sensory words produce mismatch expressions like smooth color and red touch in contrast to normal expressions like red color and smooth touch. Concerning these sensory mismatch expressions, results of three experiments are reported. Experiment 1 revealed that (i) mismatch expressions were less comprehensible than normal expressions, and that (ii) there were two patterns among mismatch expressions: the high-comprehensible mismatch expression (HighCME, e.g., smooth color) and the low-comprehensible mismatch expression (LowCME, e.g., red touch). Experiment 2 revealed that the mismatch expressions produced a significantly greater N400 amplitude than the normal expressions. Experiment 3 implied that the difference between High/LowCME was reflected in a later latency band or in a topographical difference of N400, although the statistical significance was marginal. It is argued that the processes to integrate linguistic elements (e.g., combining adjectives and nouns) are not homogeneous.

Keywords: Event-related potentials (ERPs); N400; Sensory mismatch; Comprehensibility; Selectional restriction; Nonliteral (metaphorical) interpretation

1. Introduction

In order to comprehend an expression like red color, we have to combine and integrate the adjective (red) and the noun (color) into the whole noun phrase. In this case, both red and color are words concerned with the ‘visual’ sense modality. As this expression shows, in an expression consisting of an adjective and a noun, the sense modality of each component is usually the same. In an expression like smooth color, however, the visual noun color is modified by the tactile adjective smooth. The combination of different sense modalities produces a ‘sensory mismatch expression.’ The expression resulting from this mismatch is a kind of metaphor called ‘synaesthesia’ in the traditional rhetoric. From a linguistic point of view, such a sensory mismatch is explained as a violation of selectional restriction (henceforth, abbreviated as SR). For example, a violation of SR arises in the expression handsome bride due to the inconsistency of the two features ‘+male’ and ‘+female’ (or ‘+male’). For review of SR, see Chomsky (1965) and Levin (1977) among others. In this study it is assumed that a sense modality such as visual or tactile also constitutes this kind of semantic feature. If these features are not in accord with each other, it is impossible to combine them. For example, the feature of being smooth (‘+tactile’) does not agree with the feature that the word color (‘+visual’) requires. In other words, the sensory mismatch expression smooth color violates the SR between these two words. By the same token, red touch is a mismatch expression that violates the SR between the visual adjective red and the tactile noun touch.

* Corresponding author. Fax: 92-642-2414.
E-mail address: sakamoto@lit.kyushu-u.ac.jp (T. Sakamoto).

© 2003 Elsevier Science (USA). All rights reserved.
do:10.1016/S0093-934X(02)00539-4
Examining these sensory mismatch expressions, however, one may notice that smooth color is somehow comprehensible, while red touch is almost incomprehensible. Despite the fact that both of them contain a sensory mismatch, there seems to exist a difference in comprehensibility. Previous researchers have revealed that there is some bias or tendency among the combinations of sensory terms. Collecting and analyzing 2000 examples from 12 poets, Ullmann (1951, p. 280) states that “transfers tend to mount from the lower to the higher reaches of the sensum from, the less differentiated sensations to the more differentiated ones, and not vice versa.” He proposes the following hierarchy from the lowest to the highest: Touch → Heat → Taste → Scent → Sound → Sight. This arrangement is said to reflect “a simple and rough-and-ready scheme reflecting the linguistic articulation, the ‘field’ of the sensum as it exists in the consciousness of the users of language.”

Williams (1976) states that the diachronic changes of sensory words have been carried out from lower to higher senses. For example, the word dull is listed in OED as a tactile word in 1230, as a visual word in 1430, and as an auditory word in 1475. In Japanese as well (Kusumi, 1988; Sakamoto, 1983), it is reported that an expression is easier to understand when the modifying direction is ‘upward’ (from lower to higher modalities) than when it is ‘downward’ (from higher to lower modalities). Thus it has been observed that the direction of combination affects the comprehensibility of mismatch expressions.

To summarize the above statements and observations, there are two different patterns of comprehensibility among expressions consisting of two elements to be integrated. A difference in comprehensibility can be caused by the existence/absence of a sensory mismatch. This is a linguistic (more strictly speaking, ‘semantic’) phenomenon in the sense that it results from the violation of a linguistic SR. Since the existence/absence of mismatch is crucial, this type of comprehensibility difference is privative and discrete. On the other hand, the difference in comprehensibility due to the combinational direction reflects a gradual degree that is changeable depending upon each speaker’s mental impression (i.e., smooth color is perceived to be more comprehensible than red touch). Thus, this type of comprehensibility difference becomes gradual and continuous. In order to construe a noun phrase that consists of an adjective and a noun, we have to integrate the two components. Is the same process of integration at work in both mismatch and normal expressions, and is the integration process the same among mismatch expressions with different degrees of comprehensibility?

In order to answer the above two questions, three experiments were conducted. Experiment 1 employed an off-line method to examine differences in comprehensibility among various combinations of sensory words through the introspection of subjects. Experiment 2 employed an on-line method to investigate what kind of brain activities were observed by analyzing ERPs elicited by sensory mismatch expressions. Furthermore, in Experiment 3, the ERP components were scrutinized by extending the spatio-temporal ranges of investigation.

2. Experiment 1

If the sensory mismatch results from the violation of a linguistic SR, we predicted that such mismatch expressions should be harder to comprehend than their normal counterparts. Furthermore, it has been observed that there is a difference in comprehensibility depending on the direction of combination between the modifying and modified terms. Thus the aim of this experiment was to clarify the following two points: (i) Does ‘the existence of sensory mismatch’ influence comprehensibility? (ii) Does ‘the difference in combinatory direction’ influence comprehensibility?

The stimuli phrases were produced by combining 22 modifying adjectives and 8 modified nouns.1 Half of them belonged to the visual modality, and the other half to tactile modality. The modifying words were selected from the list of high frequency (more than .014 per mill) words in Kusumi (1988). The familiarity of the modified nouns was examined using a questionnaire with a five-point scale. Averages were not different between visual and tactile nouns: $F_1(1, 22) = 1.116, p = .3023; F_2(1, 6) = .105, p = .7566$. Thus, the frequency between the two groups of modifying adjectives was the same, and the two groups of modified nouns had the same familiarity. Therefore, all the combinations should be equal in calculation. As shown in Table 1, however, the actual ratings of comprehensibility, which will be explained later, were not the same. Some examples of the phrases used as the stimuli are shown in Table 1.

A total of 176 phrases were used as the stimuli (see Appendix). These phrases consisted of 88 sensory mismatch expressions, and 88 normal expressions (without any sensory mismatch). Among the normal expressions, those which have visual and tactile nouns are respectively Conditions A and B. Among the sensory mismatch expressions, the upward and downward combinations are respectively Conditions C and D.

---

1 There are two types of adjectives in Japanese: “-i adjectives” like aka-i ‘red’ and “-na adjectives” like nameraka-na ‘smooth.’ In addition to these adjectives, the adjective list here includes some verbs inflected to modify nouns. They are comparable with the examples of present and past participles like sparkling star and polished glass. These inflected verb forms are regarded as adjectives in the sense that they have the ability to modify nouns.
Table 1  
Examples of the sensory-word term combinations and their comprehensibility ratings (standard deviation)

<table>
<thead>
<tr>
<th>Condition</th>
<th>Modifier-noun</th>
<th>Example</th>
<th>Mismatch</th>
<th>Comprehensibility ratings (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Condition A</td>
<td>Visual–visual</td>
<td>Red color</td>
<td>– (None)</td>
<td>4.416 (0.811)</td>
</tr>
<tr>
<td>Condition B</td>
<td>Tactile–tactile</td>
<td>Smooth touch</td>
<td>– (None)</td>
<td>4.144 (1.086)</td>
</tr>
<tr>
<td>Condition C</td>
<td>Tactile–visual</td>
<td>Smooth color</td>
<td>+ (Upward)</td>
<td>3.298 (1.230)</td>
</tr>
<tr>
<td>Condition D</td>
<td>Visual–tactile</td>
<td>Red touch</td>
<td>+ (Downward)</td>
<td>1.679 (0.841)</td>
</tr>
</tbody>
</table>

2.1. Method

The subjects were 57 students of Kyushu University ranging from 19 to 22 years of age. Experiment 1 was an off-line experiment employing a questionnaire with a five-point scale. All phrases were printed on sheets of paper and a five-scale measure ranging from the least comprehensible (scale 1) to the most comprehensible (scale 5) was printed beside each phrase. A subject was asked to rate the comprehensibility of these 176 phrases. A rating was performed by marking the five-scale measure to the right side of each phrase. The subjects performed this task at their own speed. In this off-line experiment, there was no time limitation in performing the task.

2.2. Results

The mean ratings of comprehensibility and standard deviation (SD) for each combination are reported in the final column of Table 1 and the ratings are plotted as in Fig. 1.

A two-way ANOVA \((2 \times 2)\) with repeated measures was performed. The factors were the existence of mismatch (mismatch vs. normal) and the difference in sense modalities in the modified noun (visual vs. tactile). Main effects were observed in both factors. For the factor of ‘mismatch’: \(F_1(1, 56) = 429.783, p < .001\); \(F_2(1, 172) = 130.189, p < .001\). For the factor of ‘modified noun’: \(F_1(1, 56) = 915.353, p < .001\); \(F_2(1, 172) = 467.71, p < .001\). The interaction of the two factors was significant: \(F_1(1, 56) = 301.516, p < .001\); \(F_2(1, 172) = 66.065, p < .001\). The post-hoc analysis revealed that both simple main effects were significant \((p < .05)\).

![Fig. 1. Comprehensibility ratings in Experiment 1.](image)

2.3. Discussion

There are two important findings concerning the comprehensibility of these expressions. First, sensory mismatch expressions have significantly lower comprehensibility than normal expressions. In other words, there is a different degree of comprehensibility due to the existence of a mismatch. Since the sensory mismatch is a category mistake resulting from the linguistic violation of SR, this result seems to be a natural consequence.

Second, it is rather easy to understand an expression when a lower modality adjective (e.g., smooth) modifies a higher modality noun (e.g., color). On the other hand, it is rather difficult to understand an expression when a higher modality adjective (e.g., red) modifies a lower modality noun (e.g., touch). Although a sensory mismatch expression results from any combination of words from two different sense modalities, there is a different degree of comprehensibility according to the combinatory direction. That is, an expression of upward direction (e.g., smooth color) can be called a high-comprehensible mismatch expression (henceforth, High-CME), while an expression of downward direction (e.g., red touch) can be called a low-comprehensible mismatch expression (henceforth, LowCME). Statistical analysis confirmed that the difference between HighCMEs and LowCMEs was significant. Thus, the second finding in Experiment 1 replicated the findings observed in the previous studies (Kusumi, 1988; Sakamoto, 1983; Ullmann, 1951; Williams, 1976).

Among the 44 HighCME phrases, note that some expressions such as soft color, cool color, and warm color are very familiar idioms (or, dead metaphors) in both English and Japanese (and perhaps in many other languages). One may argue that these kinds of idioms have already been listed in our mental lexicon, so they produced high comprehensibility. It may seem reasonable to claim that soft color produces high comprehensibility because it is an idiom. However, we have to explain why it is an idiom. One could say that it became an idiom because it is an upward combination. Thus, to say that an idiom produces high comprehensibility is the same as to say that the upward combination produces high comprehensibility.

Furthermore, whether an expression becomes an idiom (or, dead metaphor) depends on some extra-linguistic factors such as culture, custom, history, etc. For example, although yellow voice is a very common idiom
3. Experiment 2

When semantic expectancy is not satisfied, a negative ERP component around 400 ms (known as ‘N400’) is elicited. It is widely accepted that the N400 is sensitive to semantic processing in language comprehension (Brown & Hagoort, 1993; Chwilla et al., 1995; Friederici, Pfeifer, & Hahne, 1993; Kutas & Hillyard, 1980, 1983; Münte, Heinze, & Mangun, 1993; Rösler, Pütz, Friederici, & Hahne, 1993; Van Petten, 1995, among others). Now, let us consider the case of semantic incongruity in Japanese. Nakagome et al. (1999) examined whether the N400 was elicited by the violation of selectional restriction (SR) between verbs and nouns. Consider the following examples. (The question mark in the sentence initial position indicates that the sentence is semantically anomalous.)

1. a. Taro-ga ryoko-ni dekake-ta.
   Taro-NOM a journey-DAT set out-PAST
   ‘Taro set out on a journey.’

   b. ?Taro-ga jisyo-ni dekake-ta.
   Taro-NOM a dictionary-DAT set out-PAST
   ‘Taro set out on a dictionary.’

   The verb dekake-ta ‘set out’ requires a place-noun indicating where to go out, but the noun jisyo ‘dictionary’ does not satisfy this requirement. It is reported that the increase of the N400 amplitude was observed to correlate with this violation of SR.

   In our case, sensory mismatch expressions induce the failure of semantic expectancy. For example, when the term red is presented, one would expect to encounter a noun that can be modified by that color term. This expectation is not satisfied when the term touch is presented. Then, we predict that LowCMEs like red touch elicit a larger N400 than normal counterparts like smooth touch. On the other hand, HighCMEs such as smooth color may or may not elicit the N400. If smooth color elicits a larger N400 than its normal counterpart red color, it may indicate that a linguistic SR influences the processing of this expression. If it does not, it may indicate that the linguistic SR is suppressed by the high comprehensibility of the HighCME. That is, the HighCME and the LowCME could be processed differently in spite of the fact that both of them involve a mismatch. In Experiment 2, therefore, we are interested in investigating whether the N400 amplitude reflects the existence of sensory mismatch and/or direction of combination.

3.1. Method

Subjects. The subjects were ten native speakers of Japanese. They were students of Kyushu University ranging from 19 to 22 years of age with normal or corrected-to-normal vision. They were paid nominal fees for their participation.

Materials. The stimuli were the same as those used in Experiment 1. All of the 176 phrases were shown to each subject in random order with word-by-word presentation. The subjects were required to read each word and to answer orally one of the five-scale measure.

Apparatus and procedure. An AV tachistoscope system (Iwatsu Isel IS-702) was used for stimulus presentation. Each subject was seated in a comfortable chair at a distance of 2 m from the screen. All the stimuli were presented at the center of the screen. First, a fixation (a star figure) appeared for 1000 ms, followed by a 600-ms blank screen. An adjective was then presented for 700 ms. After a 600-ms blank screen, a noun was

---

2 This idiom is used to describe the high-pitched voice of some children and some young women. The ‘yellow’ seems to be associated with the bill or beak of nestlings.
presented for 600 ms. Each word subtended a visual angle of 0.6° to 3.0° horizontally. The participants were required to read the phrases carefully. The presentation order of the phrases was determined randomly for each subject.

**ERP recording.** An evoked potential measuring system was used (Nihon Kohden, MEB-5504). The electroencephalogram (EEG) activity was recorded by means of Ag/AgCl electrodes from three scalp locations (Cz, C3, and C4), based on the International 10–20 System. All the EEG electrodes were referenced to linked earlobes. The ground electrode was positioned on the forehead. Eye movements and blinks were monitored via an electrode placed beneath the left eye. Electrode impedances were kept below 5 kΩ.

**ERP analysis.** ERPs were averaged for an epoch of 2000 ms, beginning with the onset of the modified noun. A 20-ms epoch before the onset of the noun was used to compute the baseline. Trials contaminated by artifacts were excluded from the analysis. Peak negativity was measured for each condition during the latency band of N400 (300–500 ms).

3.2. Results

Grand average waveforms from Experiment 2 are shown in Fig. 2. A three-way ANOVA (2 × 2 × 3) with repeated measures was performed. The factors were (i) the existence of sensory mismatch (mismatch vs. normal), (ii) the sense modality to which the modified noun belongs (visual vs. tactile), and (iii) electrodes (C3, Cz, and C4). All three factors were within-subject factors. Concerning the N400 amplitude, we can summarize the results as follows:

(i) A significant main effect was observed concerning the existence of sensory mismatch ($F(1,9) = 9.383, p = .0135$).

(ii) The interaction between the mismatch and the electrodes was significant ($F(2,18) = 4.776, p = .0217$). The post-hoc analysis revealed that the N400 amplitude of Cz and C4 was significantly greater in conditions with a sensory mismatch than in conditions without such a mismatch. Cz: $F(1,27) = 13.565, p = .001$; C4: $F(1,27) = 2.252, p = .0041$.

(iii) A main effect of modified noun was not observed: $F(1,9) = 1.092, p = .3233$.

3.3. Discussion

From the result of (i), shown above, it became clear that sensory mismatch expressions (e.g., *smooth color* and *red touch*) produced the greater N400 amplitude than normal expressions (e.g., *red color* and *smooth touch*). In other words, Experiment 2 revealed that the N400 reflected the processing of a sensory mismatch involved in both HighCMEs and LowCMEs. The electrophysiological measure of Experiment 2 therefore verified the first finding in Experiment 1: that comprehensibility differs due to the existence of a mismatch.

The result of (ii), reported above, indicates that the increase in the N400 amplitude is remarkable around the right-center region. This may align with the scalp distribution of the typical N400, which is larger over the right hemisphere (Kutas & Van Petten, 1994; Nakagome et al., 1999; Neville, Nicol, Bars, Forster, & Garrett, 1991, among others). However, attention is necessary for this distributional pattern because the distribution of ERPs is not so precise. More detailed data would be necessary for speculation about the activated regions inside the brain, since the analysis was based on only three electrodes in this experiment.

The result reported in (iii) above indicates that the difference in the modality of the modified noun did not influence the N400 amplitude. This means that there is no difference in the N400 amplitude between HighCMEs (e.g., *smooth color*) and LowCMEs (e.g., *red touch*). The N400, therefore, does not reflect a difference in comprehensibility due to the direction of combination, which was observed in Experiment 1. Then, is it the case that there are no such brain activities reflecting the
difference in comprehensibility between HighCMEs and the LowCMEs? In order to answer this question, let us consider some possible moves. Suppose that we ignore the direction of combination, and select 'the most highly comprehensible phrases' and 'the least comprehensible phrases.' For example, suppose that we select 10 phrases each from the top and bottom among the 88 mismatch expressions. This contrast between the two extremes may induce some difference in brain activities that is somehow detectable by ERPs. However, there are two problems with this move. First, we need more than 40-times accumulation to get an average for data analysis, since the electroencephalogram (EEG) activity is very subtle. Thus, we need many more examples than we have now, although this is a technical problem that can be resolved in any case. Second and more severely, this approach is too ad hoc to explain anything. Suppose that we collect enough instances from the top and bottom of comprehensibility and then find that the two groups are correlated with some different brain activities. This result would seem to be very plausible. However, it only says that highly comprehensible phrases are highly comprehensible because they are highly comprehensible. The claim is a tautology and the discussion is circular.

Then, one of the most promising moves to answer this question is to extend the temporal and the spatial coverage areas of investigation (see Friederici, Mecklinger, Spencer, Steinhauer, & Donchin, 2001). If this extension of the research area fails to detect any difference between HighCMEs and LowCMEs, then this difference may not be psychologically real. If, however, this difference is reflected in the spatial and/or temporal aspects of ERPs, it would provide us with a piece of evidence that the High/LowCME distinction is psychologically real. To get more extensive and detailed data we planned Experiment 3 as explained in the next section.

4. Experiment 3

Employing an off-line measure, Experiment 1 revealed that there are two types of 'difference in comprehensibility': one is caused by the existence of mismatch itself, and the other is produced by the direction of combination. Employing an on-line measure, Experiment 2 revealed that the N400 reflected a difference in comprehensibility caused by the existence of mismatch. However, this experiment failed to detect the comprehensibility difference due to the combinatory direction. The natural move to get more information on this matter is to extend the spatio-temporal ranges of investigation, which would lead to the following two predictions.

(i) The brain activities corresponding to the directional difference in comprehensibility can be reflected in a different latency band from the N400. We are especially interested in the latency bands before and after the N400: an early negative component (ENC, 50–150 ms) and a late positive component (LPC, 500–900 ms). Since the brain activities associated with various linguistic phenomena are so complicated, we cannot discard the possibility that the difference in comprehensibility (i.e., between HighCMEs and LowCMEs) is correlated with various ERP components (for a very concise review of various ERP components, see Friederici, 2000; Kutas & Van Petten, 1994; and Townsend & Bever, 2001).

(ii) Relevant activities can be reflected in different regions of the brain, although the latency band may be around the N400 component. Therefore, the difference in comprehensibility due to the combinatory direction may be reflected in a difference in the activated regions of the brain. That is, the spatial distribution of N400 can be relevant to this difference. It is well known that the right and left hemispheres have different functions. Brownell, Potter, Michelow, and Gardner (1984) pointed out that right-hemisphere-damaged patients were insensitive to connotative meanings (e.g., cold has a connotative meaning of remoteness). See also Anaki, Faust, and Kravetz (1998) and Winner and Gardner (1977) for the importance of the right hemisphere to understand metaphorical expressions. In understanding very familiar and highly conventional metaphors like broken heart, however, it is reported that the left hemisphere is at work (Giora, 1999). Thus, it might be the case that a LowCME like red touch is processed in the right hemisphere as an unfamiliar metaphor, while a HighCME like smooth color is processed in the left hemisphere as a familiar metaphor. Therefore, we are interested in the topographical difference in activation between HighCMEs and LowCMEs.

4.1. Method

The subjects were 11 native speakers of Japanese. They were students of Kyushu University ranging from 19 to 22 years of age with normal or corrected-to-normal vision. They were paid nominal fees for their participation. Each phrase was shown to the subjects with word-by-word presentation. They were asked to read it silently. For every 26 to 32 trials, the subjects were asked to respond orally to the question "What was the last
phrase?” This was the procedure to maintain the concentration of the subjects.

The experimental design and stimuli were the same as those in Experiment 2, except that the number of electrodes was increased from three to seven (Fz, Cz, C3, C4, Pz, P3, and P4) in Experiment 3. The record of ERP and EOG was started 200 ms prior to the eye-fixation point. The baseline for an ERP analysis was the average of the electric potentials 100 ms prior to the presentation of the modified noun.

4.2. Results

We analyzed the average ERPs in the following three latency bands: the ENC (early negative component, 50–150 ms), the N400 component (300–500 ms), and the LPC (late positive component, 500–900 ms). A three-way ANOVA ($2 \times 2 \times 7$) with repeated measures was performed. The factors were (i) the existence of mismatch (mismatch vs. normal), (ii) the modality of the modified noun (visual vs. tactile), and (iii) electrodes. Focusing on the important aspects, we can summarize the results of this experiment in the following three points.

First, concerning the ENC, we found a marginally significant main effect of mismatch ($F(1, 10) = 3.309, p = .0989$). For all of the differences reported in this paper, we take $p < .05$ to indicate statistical significance; $.05 < p < .10$ to constitute marginal significance. This marginal effect in the ENC may suggest that the sensory mismatch was monitored in the very early stage of processing. In other words, a rather automatic process to detect semantic incongruity was brought about by this mismatch. However, note that the statistical results indicate that this is only a ‘tendency.’ So, the empirical support is weak. In this latency band, there was no effect of modified noun: $F(1, 10) = 2.214, p = .1676$. This means the difference in comprehensibility between a High- and a LowCME is not reflected in the ENC latency band.  

Second, in the N400 component, a significant main effect of mismatch was observed: $F(1, 10) = 6.122, p = .0329$. The N400 amplitude of mismatch expressions was significantly greater than that of normal expressions. However, there was no significant main effect of modified noun: $F(1, 10) = .687, p > .4264$.

Third, in the LPC, there was no significant main effect of mismatch: $F(1, 10) = 2.128, p = .1753$. In this latency band, however, we found a marginally significant main effect of modified noun: $F(1, 10) = 4.8, p = .0533$. This may suggest the possibility that the difference in comprehensibility due to the combinatory direction is reflected in a later latency band rather than the N400. Statistically, however, since this is only a tendency, the empirical basis is not strong enough to support this possibility. The LPC could be a version of SPS (syntactic positive shift) or P600, which is assumed to be elicited by various syntactic violations such as subjunction violation and agreement violation (Gunter & Friederici, 1999; Hahne & Friederici, 1999; Osterhout, Holcomb, & Swinney, 1994). If the LPC is sensitive only to syntactic violations, it is reasonable that we could not find the semantic mismatch effect.

4.3. Discussion

Considering these experimental results from the temporal point of view, we got basically the same results as Experiment 2. That is, the N400 reflects the difference between mismatch and normal expressions, while it does not reflect the difference between High- and LowCMEs (although the possibility that the LPC reflects this difference is implied). The findings of Experiment 3, thus, seem to be the simple replication of Experiment 2.

Since only the N400 component exhibited the significant main effect of mismatch, a more detailed examination of this component would supply us with more information to account for the High/LowCME distinction. When we subtract the N400 amplitude of normal expressions (e.g., red color) from that of HighCMEs (e.g., smooth color), as can be seen in Fig. 3 we find that the activation spread widely. On the other hand, when we subtract the N400 amplitude of normal expressions (e.g., cold touch) from that of LowCMEs (e.g., red touch), we notice that activation occurred only in the right hemisphere (C4 and P4).

It is not a simple matter to determine the spatial relations between the electrical activities recorded on the scalp and the neurological source of those activities. It cannot be concluded that the distribution of ERPs precisely corresponds with the functional distribution of the cerebrum. As far as there is distributional “difference” in the ERPs, however, it is expected that the difference probably has some relationship to some functional difference in the hemispheres. Thus, we conducted a planned (a priori) t-test with the hemispheric difference as a factor (right vs. non-right). This planned comparison suggested a marginally significant effect of the hemispheric difference for LowCMEs ($t_{10} = -2.186, p = .0537$), but not for HighCMEs ($t_{10} = .101, p = .9212$).  

\footnote{Although these analyses were performed, strictly speaking, they should have been justified by an interaction among the three factors, which did not reach significance. When we have a planned hypothesis to be examined a priori, however, this kind of planned (a priori) comparison is possible (e.g., Kim, 1999).}

\footnote{The characteristics of this ERP component are unclear even if this component is proved to exist statistically. It could be a so-called N100, MMN (mismatch negativity), PN (processing negativity), or ELAN (early left anterior negativity). See Friederici (2000) and Näätänen (1989) for more detail.}
smooth color increased the N400 amplitude in the whole region, while a LowCME like red touch produced activation only in the right hemisphere. In other words, regardless of the difference in comprehensibility due to the combinatory direction, both High- and LowCMEs activated the right hemisphere. In addition to activation of the right hemisphere, HighCMEs activated the left-central region, too. Note, however, that the statistical analysis does not fully support the discussion of this topological difference, since the significance level was marginal.

In sum, the existence/absence of mismatch produces a privative and discrete difference in comprehensibility, which has a clear N400 effect. Thus, the process to integrate the modifying adjective and the modified noun could be different between mismatch and normal expressions. On the other hand, the difference in combinatory direction results in a gradual and continuous comprehensibility, which exhibits only marginal effects on brain activities. The integrating process to construe the noun phrase could be the same among the mismatch expressions, although the degree of comprehensibility is gradually different.

5. General discussion

We can summarize the results of the three experiments concerning the two factors: the existence of sensory mismatch and the modalities of the modified noun.

The results of Experiment 1 revealed that there was a difference in comprehensibility depending on the existence of sensory mismatch and on the direction of combination (upward vs. downward). The on-line analysis in Experiment 2 showed that the N400 amplitude reflected the existence of sensory mismatch. The results of Experiment 3 implied two possibilities to account for the difference due to combinatory direction. One is that this difference is reflected in some later latency band such as the LPC, while the other is that this difference is reflected in the topographical distribution.

Examining some previous studies, we consider three aspects of the issue in the following sections. (i) The characteristics of the N400 with relation to the mismatch, (ii) "temporal" aspects of non-literal meanings, and (iii) "spatial" aspects of non-literal meanings.

5.1. The N400 due to violation of linguistic restriction

Both Experiments 2 and 3 have indicated that the N400 amplitude increased when the target phrase involved a sensory mismatch. Let us briefly consider the characteristics of this ERP component. Rösler et al. (1993) examined some ERP components with respect to semantic and syntactic violations with following examples. (The asterisk and the question mark in the sentence initial position indicate respectively syntactic or semantic anomalies.)

(2)

a. Der Päsident wurde begrüßt. (The president is being greeted.)
b. *Der Lehrer wurde gefallen. (The teacher is being fallen.)

(3)

a. Der Clown hat gelacht. (The clown has laughed.)
b. *Der Dichter hat gegangen. (The poet has gone.)

(4)

a. Das Paket wurde geliefert. (The parcel is being delivered.)
b. ?Der Honig wurde ermodet. (The honey is being murdered.)

(5)

a. Die Kerze hat gebrannt. (The candle has burned.)
b. ?Der Ball hat geträumt. (The ball has dreamed.)

The syntactic violation in (2b) and (3b) is caused by the incorrect choice of the auxiliary wurde ‘is being’ and hat ‘has.’ It was reported that these syntactic anomalies evoked an N400-like component around the anterior region of the left hemisphere. The semantic violation in (4b) and (5b) is caused by the inappropriate selection of inanimate subjects for animate verbs. It is reported that these semantic anomalies evoked a widespread N400-like component over posterior temporal areas.

Rösler et al. (1993) found that both syntactic and semantic violations elicited a negativity around 400 ms. Thus, it may be the case that the ‘processing of mismatch’ in general elicits the negative component around
400 ms. It has been pointed out that the N400 reflects the integration process of two linguistic elements (Brown & Hagoort, 1993; Chwilla et al., 1995; Holcomb, 1993; Kutas & Hillyard, 1983; Rugg, 1990, among others). The N400 elicited in the processing of sensory mismatch expressions seems to reflect this kind of integration process as a way of dealing with incongruity due to the violation of linguistic restriction.

5.2. Non-literal (metaphorical) meanings and temporal aspects of ERPs

Pynte, Besson, Robichon, and Poli (1996) observed that non-literal (metaphorical) sentences (e.g., ‘Those fighters are lions.’) elicited significantly larger N400 amplitude than literal sentences (e.g., ‘Those animals are lions.’) \( (p < .009) \). However, they failed to detect an effect of nonliteral interpretation in a later latency band such as the LPC. Furthermore, the N400 amplitude was not different between familiar metaphors (e.g., ‘Those fighters are lions.’) and unfamiliar metaphors (e.g., ‘Those apprentices are lions.’) \( (p > .10) \). The results of their experiments revealed the following two points: (i) The effect of nonliteral interpretation is reflected in the N400 component, but not in the LPC. (ii) The N400 does not reflect a difference of familiarity in nonliteral meanings. Based on these experimental findings, Pynte et al. (1996) rejected the following two hypotheses about the processing of nonliteral meaning: the hierarchical and parallel hypotheses. The hierarchical hypothesis states that nonliteral meaning is accessed after literal meaning is rejected, predicting that a later ERP component such as the LPC reflects nonliteral interpretation. The parallel hypothesis assumes both literal and nonliteral meanings are processed at the same time, predicting that the amplitude of an ERP component such as N400 further increases.

These findings by Pynte et al. seem to be similar to our findings. In our experimental settings, the hierarchical hypothesis predicts that a component later than the N400 latency band would reflect the directional difference. In our Experiment 3, it was implied that this difference would be reflected in the LPC. That is, it would be possible for this difference to correspond to the difference of difficulty in searching for non-literal meaning. However, this possibility was not sufficiently supported by the statistical analysis. Thus, our experimental results did not fully support the hierarchical hypothesis.

On the other hand, the parallel hypothesis assumes that both mismatch and combinatory direction influence the ERP at the same latency band. In Experiments 2 and 3, as shown in Table 2, there was no latency band in which both the mismatch and combinatory direction were significant. That is, it was not the case that both factors were reflected in the same ERP component at the same time. This result does not support the prediction by the parallel hypothesis.

To summarize, both Pynte et al.’s and our experimental results support neither the hierarchical nor the parallel hypothesis. Pynte et al. thus argued for a third hypothesis, the context-dependent hypothesis, which states that a context becomes the clue to discovering nonliteral meaning before accessing literal meaning. However, we cannot evaluate the validity of this hypothesis because no context was supplied in our experiments. Furthermore, we assume that it is not sufficient to examine hypotheses which are based solely on the temporal aspects of processing. In what follows, therefore, we examine a hypothesis in which not only the temporal but also the spatial condition is taken into account.

5.3. Hemisphere asymmetry and distributional difference of ERPs

A dual-coding theory assumes two processing systems (verbal and imaginal) for a concrete word (e.g., TABLE) and just one processing system (verbal) for an abstract word (e.g., JUSTICE). The verbal system is assumed to be linked with the left hemisphere, and the imaginal system is linked with the right hemisphere (for a review of dual-coding theory, see Paivio, 1991). Kounios and Holcomb (1994) showed that there was a definite difference in topographical distribution of ERPs in processing concrete and abstract words. The concrete words

<table>
<thead>
<tr>
<th>Experiment 1</th>
<th>Experiment 2</th>
<th>Experiment 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Off-line</td>
<td>On-line</td>
<td></td>
</tr>
<tr>
<td>A: mismatch</td>
<td>***</td>
<td>+</td>
</tr>
<tr>
<td>B: modified noun</td>
<td>*</td>
<td>n.s.</td>
</tr>
<tr>
<td>AB interaction</td>
<td>***</td>
<td>n.s.</td>
</tr>
</tbody>
</table>

n.s. (non-significant).

* \( p < .05 \).

*** \( p < .001 \).

* \( p < .10 \).
increased the N400 amplitude in both hemispheres. In this case, both verbal and imaginal processes seemed to be at work. On the other hand, the abstract words activated only the left hemisphere, relating to verbal processing.

In our Experiment 3, the results implied that the topographical difference of ERP could reflect a difference in comprehensibility due to the combinatorial direction. This difference may correspond to the difference between the literal processing system in the left hemisphere and the non-literary (or, metaphorical, imaginal) processing system in the right hemisphere. Because a LowCME such as red touch is not a normal but a mismatch expression, it might be that the normal verbal process for normal expressions like smooth touch in the left hemisphere could not cope with it. That is why a LowCME is processed only by the non-literary system in the right hemisphere. On the other hand, because a HighCME such as smooth color can be processed as a typical verbal expression like a familiar idiom, it could be the case that the normal verbal process in the left hemisphere worked actively. Furthermore, because a HighCME is an expression which does contain a mismatch, it is possible that the non-literary process was working in the right hemisphere, too. In processing the HighCMEs, both the literal and non-literary processes may have been working in both the right and left hemispheres. On the other hand, only the non-literary process may have been working in the LowCMEs, which does not allow a typical literal processing in the left hemisphere. So, it may be possible that HighCMEs and LowCMEs are processed in different places at the same time. Note again, however, that the statistical support was not enough to insist on this possibility.

6. Concluding remarks

The findings in this research are especially remarkable in two points. One is that the N400 amplitude was increased by the existence of a sensory mismatch. This finding is consistent in both on-line experiments. Thus, it would be plausible to conclude that the N400 is sensitive to violation of linguistic (semantic) selection restrictions. The other finding is that the difference in comprehensibility due to the combinatorial direction has only a marginally significant effect. The possibility was implied that this kind of difference is reflected in different latency bands like the LPC and/or in the distributional difference in hemispheres. However, this possibility was not sufficiently supported by the statistical analysis.

From the first finding, we conclude that the violation of linguistic (semantic) restriction, which induces a privative and discrete difference in comprehensibility, exhibits clearly significant effects in brain activity. From the second finding, on the other hand, we conclude that a difference in combinatorial direction, which induces a gradual and continuous comprehensibility, exhibits only marginal effects in brain activities. In constructions consisting of modifying adjective and modified noun, we have to integrate the two components to construe the whole meaning. These findings revealed that there are various patterns of integration processes by inspecting electrophysiological correlations. The processes of integration would be discretely different between mismatch and normal expressions. On the other hand, it is implied that the difference of integration processes between High- and LowCMEs could be marginally heterogeneous in terms of spatio-temporal distribution of brain activities.

Appendix. List of materials

<table>
<thead>
<tr>
<th>Visual modifiers</th>
<th>Tactile modifiers</th>
</tr>
</thead>
<tbody>
<tr>
<td>akai</td>
<td>namerakana</td>
</tr>
<tr>
<td>kuroi</td>
<td>arai</td>
</tr>
<tr>
<td>shiroi</td>
<td>katai</td>
</tr>
<tr>
<td>aoi</td>
<td>yawarakai</td>
</tr>
<tr>
<td>akaru</td>
<td>tsunetai</td>
</tr>
<tr>
<td>kurai</td>
<td>atatakai</td>
</tr>
<tr>
<td>kagayakinoaru</td>
<td>kawaita</td>
</tr>
<tr>
<td>sunda</td>
<td>shimetta</td>
</tr>
<tr>
<td>toomeina</td>
<td>surudoi</td>
</tr>
<tr>
<td>nigotta</td>
<td>nibui</td>
</tr>
<tr>
<td>azayakana</td>
<td>nebakkoi</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Visual nouns</th>
<th>Tactile nouns</th>
</tr>
</thead>
<tbody>
<tr>
<td>iro</td>
<td>tezawari</td>
</tr>
<tr>
<td>iroai</td>
<td>hadazawari</td>
</tr>
<tr>
<td>shikicho</td>
<td>kanshoku</td>
</tr>
<tr>
<td>shikisai</td>
<td>sawarigokochi</td>
</tr>
</tbody>
</table>
References


