



OPEN ACCESS

EDITED BY

Tetsuo Kida,
Institute for Developmental Research, Japan

REVIEWED BY

Atsuko Gunji,
Yokohama National University, Japan
Yukiyasu Yaguchi,
Seitoku University, Japan

*CORRESPONDENCE

Kaori Sasaki
✉ kaori.s@iuhw.ac.jp

†PRESENT ADDRESS

Seiichi Kadowaki,
Department of Otolaryngology, Faculty of
Medicine, University of Tsukuba, Tsukuba,
Japan,
Department of Otolaryngology, Tokyo-Kita
Medical Center, Tokyo, Japan

RECEIVED 19 April 2023

ACCEPTED 25 September 2023

PUBLISHED 11 October 2023

CITATION

Sasaki K, Kadowaki S, Iwasaki J, Pijanowska M
and Okamoto H (2023) Cognitive neural
responses in the semantic comprehension of
sound symbolic words and pseudowords.
Front. Hum. Neurosci. 17:1208572.
doi: 10.3389/fnhum.2023.1208572

COPYRIGHT

© 2023 Sasaki, Kadowaki, Iwasaki, Pijanowska
and Okamoto. This is an open-access article
distributed under the terms of the [Creative
Commons Attribution License \(CC BY\)](#). The
use, distribution or reproduction in other
forums is permitted, provided the original
author(s) and the copyright owner(s) are
credited and that the original publication in this
journal is cited, in accordance with accepted
academic practice. No use, distribution or
reproduction is permitted which does not
comply with these terms.

Cognitive neural responses in the semantic comprehension of sound symbolic words and pseudowords

Kaori Sasaki^{1*}, Seiichi Kadowaki^{2†}, Junya Iwasaki¹,
Marta Pijanowska^{3,4} and Hidehiko Okamoto²

¹Department of Speech and Hearing Sciences, International University of Health and Welfare, Narita, Japan, ²Graduate School of Medicine, International University of Health and Welfare, Narita, Japan, ³Office of Medical Education, International University of Health and Welfare, School of Medicine, Narita, Japan, ⁴Graduate School of Humanities and Sociology, University of Tokyo, Tokyo, Japan

Introduction: Sound symbolism is the phenomenon of sounds having non-arbitrary meaning, and it has been demonstrated that pseudowords with sound symbolic elements have similar meaning to lexical words. It is unclear how the impression given by the sound symbolic elements is semantically processed, in contrast to lexical words with definite meanings. In event-related potential (ERP) studies, phonological mapping negativity (PMN) and N400 are often used as measures of phonological and semantic processing, respectively. Therefore, in this study, we analyze PMN and N400 to clarify the differences between existing sound symbolic words (onomatopoeia or ideophones) and pseudowords in terms of semantic and phonological processing.

Methods: An existing sound symbolic word and pseudowords were presented as an auditory stimulus in combination with a picture of an event, and PMN and N400 were measured while the subjects determined whether the sound stimuli and pictures match or mismatch.

Results: In both the existing word and pseudoword tasks, the amplitude of PMN and N400 increased when the picture of an event and the speech sound did not match. Additionally, compared to the existing words, the pseudowords elicited a greater amplitude for PMN and N400. In addition, PMN latency was delayed in the mismatch condition relative to the match condition for both existing sound symbolic words and pseudowords.

Discussion: We concluded that established sound symbolic words and sound symbolic pseudowords undergo similar semantic processing. This finding suggests that sound symbolism pseudowords are not judged on a simple impression level (e.g., spiky/round) or activated by other words with similar spellings (phonological structures) in the lexicon, but are judged on a similar contextual basis as actual words.

KEYWORDS

semantic comprehension, sound symbolism, onomatopoeia, event-related potential, lexical semantic processing

Abbreviations: EEG, electroencephalography; ERPs, event-related potentials; PMN, phonological mapping negativity.

1. Introduction

The arbitrary relationship between form and meaning is one of the essential characteristics of language (Hockett, 1960). Language is said to be arbitrary as there is usually no special reason why a specific form (sound or shape) is used to express a certain meaning. For example, different words, such as “apple” in English and “リンゴ (ringo)” in Japanese, are used to convey the same idea. An exception to this rule, however, is onomatopoeia. Onomatopoeia is the imitation of an actual sound that has become a standard lexical item (e.g., “bow wow” for a dog’s bark). As such, onomatopoeia is considered a lexeme with a non-arbitrary connection between form and meaning (Pinker, 1999). This kind of non-arbitrary association of sound and meaning is known as sound symbolism. Sound symbolism is defined as “the direct link between sound and meaning” (Hinton et al., 1994).

Sound symbolism has been demonstrated in experiments in which participants consistently associate certain types of phonemes with specific shapes. The phoneme is the smallest unit of sound used in the word. For instance, plosives such as /t/ and /k/ are often perceived as referring to a straight line or sharp-edged figure, and nasals like /m/, /n/ as referring to a curved line or rounded figure (Köhler, 1947; Ramachandran and Hubbard, 2001). It is believed that the sound symbolism of some phonemes and phoneme clusters is connected to the structural characteristics of the brain or the mechanism by which information is processed in the brain (Sidhu and Pexman, 2018).

While many such sound symbolic elements are consistently associated with certain meanings, they are not usually considered a part of the mental lexicon. The mental lexicon is a set of words that humans retain in their brains, the idea behind a mental lexicon is that the vast majority of information related to a word is stored in long-term memory as a concept (Aitchison, 2012). Usually, using this mental lexicon is necessary to comprehend a word’s meaning (Stille et al., 2020). Words are encodings of a specific object, such as an event, concept, or thing, with a linguistic label (Abe et al., 1994). Moreover, each word has a distinct pronunciation, spelling, meaning, and syntactic category, such as noun or verb, which is stored in a mental lexicon. In contrast, sound symbolic elements are not associated with any syntactic category and their semantic content is much vaguer compared to standard lexical items—they only provide impressions, such as round, soft, hard, or sharp (Ramachandran and Hubbard, 2001).

On the other hand, established sound symbolic words, such as onomatopoeia, can be considered full-fledged lexical items. The inventory of such words in Indo-European languages is mostly limited to onomatopoeia that express the sounds made by animals (ex., quack-quack) or sounds made while using certain objects (boom or swish). However, many languages outside of the Indo-European family use sound symbolic words known as ideophones that can describe manners, states and emotions (Osaka, 1999; Akita, 2017). In this study, we focus on Japanese, in which a great variety of such sound symbolic words is frequently used. Japanese onomatopoeia and ideophones both usually incorporate sound symbolic elements with consistent semantic associations (ex./k/: gives the image of hitting a hard surface,/g/: gives the impression of increased strength/weight) (Hamano, 2014), and use identical morphological patterns (often a reduplication of 2-mora elements,

as in/fuwa-fuwa/) (Tamori and Schourup, 1999). Onomatopoeia and ideophones are established items of Japanese mental lexicon and in this study, we use the term sound symbolic words to refer to both of these categories.

Phonological mapping negativity (PMN) and the N400 response of event-related potentials (ERPs) have been used as indicators in studies of the processing mechanisms of words and pseudowords. A negative evoked brain response connected to language processing, known as PMN, which has an earlier latency than N400, manifests around between 250 and 300 ms (Desroches et al., 2009). According to van den Brink et al. (2001), PMN represents the stage at which, in relation to the lexical selection process, word-form information resulting from an initial phonological analysis and content information derived from the context interact. However, Newman et al. (2003), and Newman and Connolly (2009) discovered that PMN occurs during the preliminary stage of language processing and reflects phonological processing. Thus, some reports of PMN indicate that it is involved in higher-level language processing, while others indicate that it is only involved in phonological processing, which occurs earlier. The interpretation of PMN depends on whether the task focuses solely on the phonological aspect of the word or on determining if the word is an appropriate choice based on higher-level linguistic characteristics.

N400 is a significant negative wave that peaks around 400 ms after the stimulus is presented (Kutas and Federmeier, 2011). The amplitude of N400 has been proven to increase for sentences with mismatched meanings (e.g., “I like coffee with cream and socks”) (Kutas and Hillyard, 1980) and is considered to be a brain potential response associated with semantic processing. Some studies of N400 interpret it as reflecting the process of retrieving words from the mental lexicon, whereas others interpret it as reflecting higher-order processes, like context specific semantic processing of vocabulary. In the former interpretation, Petten and Kutas (1990) showed that the amplitude of N400 increases for low frequency words compared to high frequency words, and Kutas and Federmeier (2000) stated that the N400 reflects the activation of vocabulary stored in long-term memory. These studies indicate that N400 may reflect access to the mental lexicon. On the other hand, Brown and Hagoort (1993) found that N400 reflects not only the process of accessing the mental lexicon but also the higher order semantic processing, as N400 is smaller when the priming effect is involved. Hagoort et al. (2004) also reported that both accessing the mental lexicon and integration into semantic knowledge happened at the same time, beginning 300 ms after language presentation. They reported that N400 is caused by the top-down influence of semantic information. In previous studies, the interpretation of PMN and N400 depended on the nature of the task and stimulus words. In their report of N400 in sound symbolic pseudowords, Deacon et al. (2004) showed that non-words derived from existing words show the same level of semantic activation as existing words when participants listen to them. On the other hand, Westbury (2005), in a study using sound symbolic pseudowords, reported that the lexical access stage, which is a preliminary stage of semantic processing, is activated, indicating that the effect of sound symbols is pre-semantic.

In past studies, the interpretation of PMN and N400 was dependent on the nature of the task and stimulus words. Moreover, previous studies on the sound symbolism of pseudowords mostly

involved judgment tasks that use simple pictures of round or angular shapes with no deeper meaning (Kovic et al., 2010; Asano et al., 2015; Sučević et al., 2015). In contrast, the task setting in this study was designed to reflect semantic processing by asking the participants to determine whether the depicted situation or object and matched or mismatched the sound stimuli (existing sound symbolic words or pseudowords). In the sound symbolic word task, there is always a correct word for the situation. On the other hand, the sound-symbolic pseudoword task is a non-word that does not correspond to any of the tasks. Sound symbolic pseudowords are non-existing words created for this study and, unlike established onomatopoeia and ideophones, cannot be considered part of the mental lexicon. Based on previous reports, it is possible that the processing of sound symbolic pseudowords also involves semantic processing; however, it is not clear whether the process is similar to that of established lexemes contained in the mental lexicon.

The goal of this study is to investigate the differences between the semantic processing of established sound symbolic words and pseudowords. Based on the results of previous studies, we suppose that the amplitude of both the PMN and the N400 will be larger in the mismatch condition than in the match condition when assessment of whether the language is consistent with more complex situations requires semantic processing. If so, the PMN appears to reflect neural processing that includes not only phonological processing, which is the precursor to language processing, but also linguistic processing (e.g., word selection). The phonological structure of sound symbolic pseudowords used the common patterns observed in existing Japanese sound symbolic words, and it included sound symbolic phonemes typical in Japanese. The phonological structure of sound symbolic pseudowords may provoke semantic neural processing. Therefore, it would support the hypothesis that sound symbolic pseudowords are semantically processed in the same way as the words if the amplitude of both the PMN and the N400 increased more in the mismatch condition than in the match condition in the sound symbolic pseudoword task, similarly to the sound symbolic words.

In the present study, we aim to clarify the processes reflected by PMN and N400 and to identify the differences in the semantic processing for existing sound symbolic words and sound symbolic pseudowords. The results obtained will give us some clues regarding the neural process of phonological and linguistic features of sound symbolism.

2. Materials and methods

Event-related potentials (ERPs) were measured and analyzed while sound symbolic word and pseudoword comprehension tasks were carried out in order to investigate the specifics of neural processing during lexical processing when a picture of an event and a speech sound are consistent (match) or inconsistent (mismatch).

2.1. Participants

A total of 30 healthy university students, ranging in age from 19 to 22, took part in the study. Participants were paid

volunteers recruited at the International University of Health and Welfare. They were all native Japanese speakers and had no medical or mental health issues in the past. They had normal or corrected-to-normal vision, and they had normal hearing. The International University of Health and Welfare's Ethical Review Committee gave its approval to this study. The experiment was explained to participants orally and in writing, and they voluntarily signed an informed consent form. This study was conducted in accordance with the ethical principles of the Declaration of Helsinki. Three participants were excluded from the study because their correct response rate to the task was less than 80%, and seven participants were excluded because the difference between their correct response rate to the sound symbolic word task and the sound symbolic pseudo-word task was more than 5%. In addition, two had many artifacts (less than 80% valid epochs) and one had incomplete task results. We used data from a total of 17 subjects for the final analysis.

2.2. Stimulus

In both sound symbolic word and pseudoword tasks, pictures and sounds were presented. In both task types, match conditions—where the appropriate sound for the picture was presented—and mismatch conditions—where a sound unrelated to the picture was presented—were used.

There were 160 stimulus word sounds, 80 each of sound symbolic words and pseudowords. The Japanese Onomatopoeia Dictionary (Ono, 2007) was used to choose existing sound symbolic words. The most prevalent type, the two-mora reduplication sound symbolic word (e.g., /fuwafuwa/, meaning “soft and puffy”) (Tamori and Schourup, 1999), was used as the sound stimulus. We chose words that express various receptive senses like touch and hearing, and the meaning is relatively clear and not context-dependent. Since sound symbolic words used in Japanese to express feelings often change their meaning depending on the situation (like/moyamoya/: thoughts and memories to be fuzzy, but also the feeling of unease or pent-up lust, depending on context), we limited our stimuli to sound symbolic words that express clear sensory states or manners. Established sound symbolic words that were used as stimulus words were modified to create sound symbolic pseudowords. The first and third syllables were not altered during creation; the second and fourth syllable were. Additionally, in keeping with Hamano's (2014) research, we changed the second syllable to a sound symbolic phoneme that elicits the same semantic impression, for example, roundness or sharpness, to preserve the word's overall semantic characteristics (e.g., /fuwafuwa/versus/fuhafuha/) ([Supplementary Data Sheet 1](#)).

A Text to speech program (Azure, Text to speech, Microsoft Inc., Redmond, WA, USA) was used to synthesize the voice of the stimulus words. All stimulus words were controlled using version: 1.0, voice type: male voice, prosody rate: 50%, pitch: 0%, used language: Japanese. We manipulated the reading speed (between 0.5 and 1.5) to fit the length of a sample to around 500 ms. In order to use the generated speech samples as stimuli for the task, the length of the speech sample was then adjusted to exactly

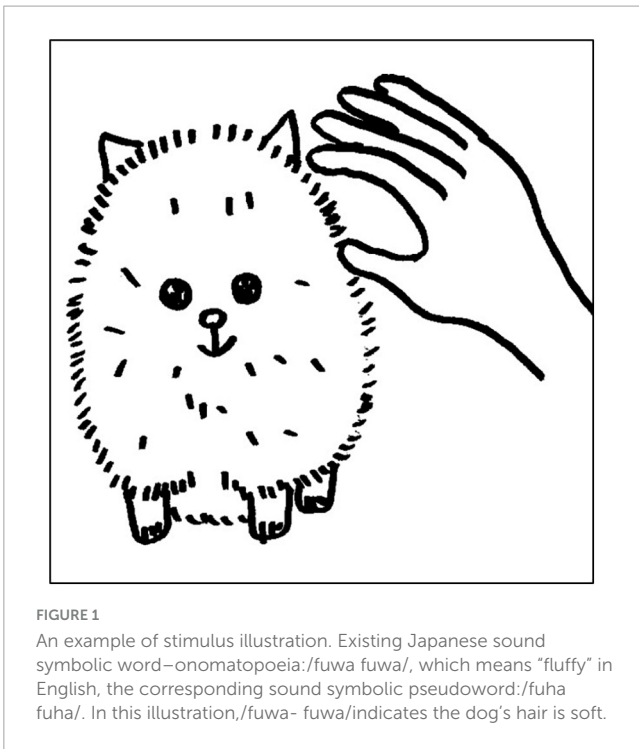


FIGURE 1

An example of stimulus illustration. Existing Japanese sound symbolic word–onomatopoeia:/fuwa fuwa/, which means “fluffy” in English, the corresponding sound symbolic pseudoword:/fuha fuha/. In this illustration,/fuwa- fuwa/indicates the dog’s hair is soft.

500 ms in Audacity 3.2, an open-source audio file editing program.¹ Fade in/out effects were added to the first and last 10 ms of each speech sample.

Both the sound symbolic word and pseudoword tasks made use of the same stimulus pictures. The meaning of established lexemes was verified in a Japanese onomatopoeia dictionary, and corresponding black-and-white pictures were created to act as the stimulus pictures (Figure 1).

There were 160 conventional sound symbolic word tasks, 80 of which were in the match condition and contained the appropriate stimulus word and picture combination. The other 80 tasks were in the mismatch condition and contained the inappropriate stimulus word and picture combination. The sound symbolic pseudoword tasks were created by replacing the existing word with a corresponding pseudoword with alternate phonemes for both the match and mismatch conditions. Consequently, there were 160 sound symbolic pseudoword tasks, 80 of which were in the match condition and 80 in the mismatch condition. Both sound-symbolic words and sound symbolic pseudowords tasks are presented twice for 40 words, for a total of 80 tasks. All of the auditory and visual stimuli were the same in the match and mismatch conditions but in different combination. The sound symbolic word and pseudoword tasks’ match and mismatch conditions added up to 320 tasks in total. In the sound symbolic word and sound symbolic pseudoword tasks, the stimulus words were presented randomly in each task.

2.3. Procedure

Participants in the experiment were seated in a soundproofed, electrically shielded room. While the participants were deciding

whether the picture on the display matched the sound coming from their earphones, their brain waves were monitored. Button-pushing tasks were given to participants to ensure they were paying close attention during the experiment. The participants were monitored while the tasks were carried out to make sure the participant was paying attention. Participants practiced on several words in both the sound symbolic task and pseudoword task prior to the actual experiment. Participants began the actual experiment when they understood the content of the task.

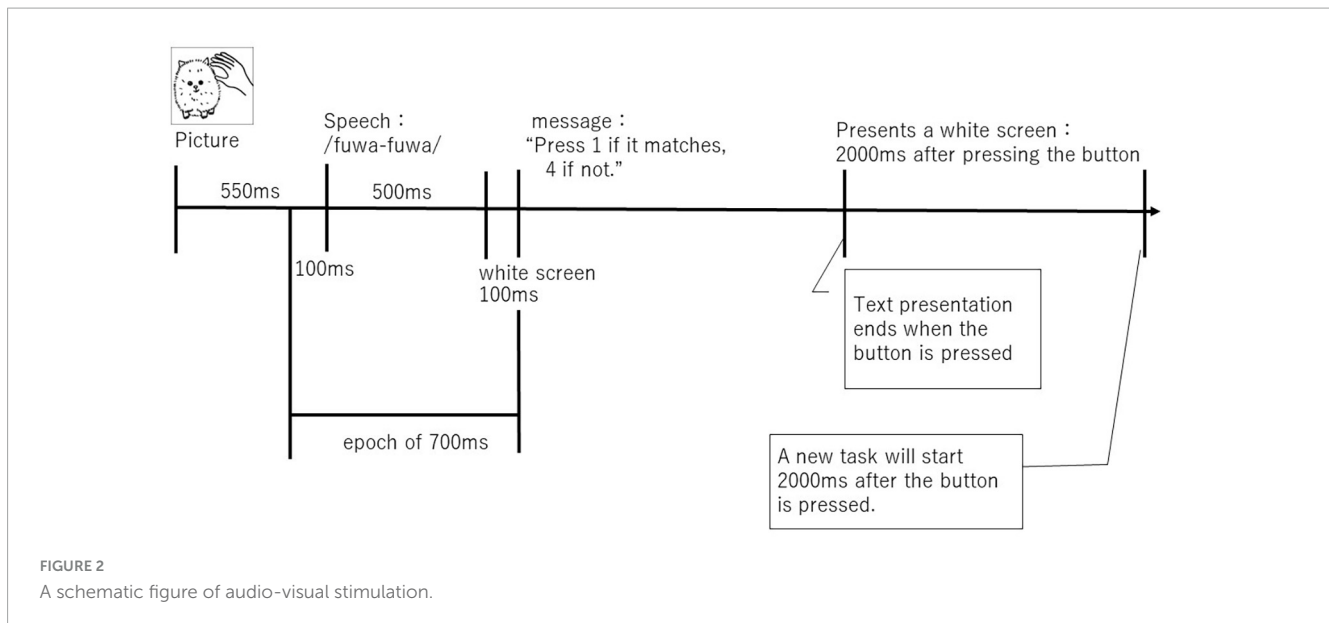
The experimental design is schematically represented in Figure 2. Stimulus pictures were shown on a 23.0” LCD screen with a resolution of 629 × 629 pixels (FlexScan EV2316, EIZO Inc., CA, USA, JP, resolution: 1,920 × 1,080), Subject-to-screen distance: range 50–85 cm, average 70 cm, and audio was played through ER-3A insert earphones (Etymotic Research Inc., IL, USA). The experiment started with the display of a stimulus picture. Next, 550 ms after the stimulus picture was displayed, a stimulus word sound (sound symbolic words or pseudowords) was presented for 500 ms. After the audio presentation, a white screen was displayed for 100 ms. Following this, a message prompting the user to interact (“Press 1 for a match, 4 for a mismatch”) was displayed in MS Gothic font size 80. Participants immediately responded by pressing the 1 or 4 key on the keyboard. The message disappeared once the participant chose a response. The next task started after 2,000 ms of a white screen. The Medical Try System Multi Trigger System Ver. 2 (MTS0410, Medical Try System, Co., Ltd., Japan) was used to control the task presentation, and the match and mismatch conditions were distributed at random.

During the experiment, participants were told to avoid blinking, swallowing saliva, or unnecessary movement. In addition, the sound symbolic word and sound symbolic pseudoword tasks were counterbalanced by randomly administering the sound symbolic word and sound symbolic pseudoword tasks to each subject. A break was always included between the pseudoword task segment and the existing word task segment. Additionally, breaks were also taken during the segments upon the participant’s request. The experiment took approximately 30–40 min to complete.

2.4. Acquiring EEG data

A unipolar recording of EEG signals was made at the Cz position. Because PMN and N400 are greatest at the parietal site (Dumay et al., 2001; D’Arcy et al., 2004; Kutas and Federmeier, 2011), PMN and N400 were measured at Cz in this experiment. Electrodes placed on the mastoids served as a reference, and the grounding electrode was positioned on the forehead. EEG signals were captured on a Neurofax EEG-1200 system (Nihon Koden, Co., Ltd., Japan). Before starting the task, it was determined that the contact impedance between the skin and electrode was less than 10 kΩ. The signal was converted to a digital signal at 1 kHz after being bandpass filtered between 0.3 and 30 Hz. The data were divided into epochs of 700 ms, after the speech stimulus presented at 100 ms. The baseline was set based on the 100 ms prior to the presentation of the speech stimulus. Reactions associated with blinking, eye movements, etc. were rejected offline based on electrooculography. Upon visual inspection, epochs with excessive noise were removed from the analysis. The total number of data points in each condition averaged 69.7 epochs (range: 79–65).

¹ <https://www.audacityteam.org>



2.5. Statistical analysis

In this study, PMN and N400 components are of particular interest. Therefore, the maximum amplitude between 250 and 350 ms after the sound onset was defined as PMN amplitude (Lee et al., 2012) and its latency was defined as PMN latency. The average amplitude between 350 and 500 ms after the sound onset was defined as N400 amplitude (D'Arcy et al., 2004; Asano et al., 2015; Manfredi et al., 2017). We analyzed PMN and N400 using Matlab R2022b (MathWorks, Inc., USA). For the existing word and pseudoword tasks, as well as for each of the amplitudes and emergent latencies, PMN was subjected to a two-way ANOVA (two factors: existing word task/pseudoword task x match/mismatch). The N400 amplitude was similarly analyzed by a two-way ANOVA using sound type (existing word vs. pseudoword) and audio-visual matching (match vs. mismatch) as factors. The statistical significance for all two-way ANOVA was expressed as a p -value of less than 0.05. Effect size was calculated using partial eta squared (η_p^2) and generalized eta squared (η_G^2) (Bakeman, 2005).

In order to more clearly analyze the differences between sound symbolic words and pseudowords, planned comparisons were conducted using paired t -tests between matching and mismatching conditions in each existing word and pseudoword task, as well as between the existing word and pseudoword task in each matching and mismatching condition. Cohen's d was used to examine effect size, with 0.2 small, 0.5 medium, and 0.8 large (Cohen, 1988). We used IBM SPSS Statistics Version 29 for Windows (IBM Corp., Armonk, NY, USA).

3. Results

3.1. Behavioral results

The average score for the sound symbolic word task for the participants was 151, and the average score percentage was 95%. On

the other hand, the mean score for the sound-symbolic pseudoword task was 149, and the mean percentage score was 93%.

3.2. Electrophysiological results

Figure 3 displays the mean Cz brain wave reading for the match and mismatch conditions of established sound symbolic words and pseudowords from 100 ms prior to speech presentation to 600 ms following presentation.

A significant main effect of sound type (existing words vs. pseudoword) was observed with the maximum amplitude in the PMN (250–350 ms) component of the ERP [$F_{(1, 16)} = 11.89$, $p = 0.003$, $\eta_p^2 = 0.426$, $\eta_G^2 = 0.030$] (Supplementary Data Sheet 2). The main effect of audio-visual matching (match vs. mismatch) was also observed [$F_{(1, 16)} = 10.07$, $p = 0.006$, $\eta_p^2 = 0.386$, $\eta_G^2 = 0.030$]. No significant interaction was found between sound type and audio-visual matching [$F_{(1, 16)} = 0.21$, $p = 0.653$, $\eta_p^2 = 0.0013$, $\eta_G^2 < 0.001$]. A significant main effect of audio-visual matching was observed for the latency at which the PMN peak appeared [$F_{(1, 16)} = 19.46$, $p < 0.001$, $\eta_p^2 = 0.549$, $\eta_G^2 = 0.085$] (Supplementary Data Sheet 3). There was no significant difference in the sound type [$F_{(1, 16)} = 0.03$, $p = 0.875$, $\eta_p^2 = 0.002$, $\eta_G^2 < 0.001$] and no significant interaction between sound type and audio-visual matching [$F_{(1, 16)} = 0.42$, $p = 0.529$, $\eta_p^2 = 0.025$, $\eta_G^2 = 0.002$]. The mean amplitude in the N400 range (350–500 ms) showed significant main effects for sound type [$F_{(1, 16)} = 16.24$, $p = 0.001$, $\eta_p^2 = 0.504$, $\eta_G^2 = 0.064$] and audio-visual matching [$F_{(1, 16)} = 36.68$, $p < 0.001$, $\eta_p^2 = 0.696$, $\eta_G^2 = 0.070$] (Supplementary Data Sheet 4). There was no significant interaction between sound type and audiovisual agreement [$F_{(1, 16)} = 0.84$, $p = 0.374$, $\eta_p^2 = 0.050$, $\eta_G^2 = 0.002$]. The amplitudes of PMN and N400 showed significant differences for sound type and audio-visual matching, but no significant interaction between them. The PMN latencies showed a significant difference for audio-visual matching, but no significant difference for sound type and no significant interaction between them.

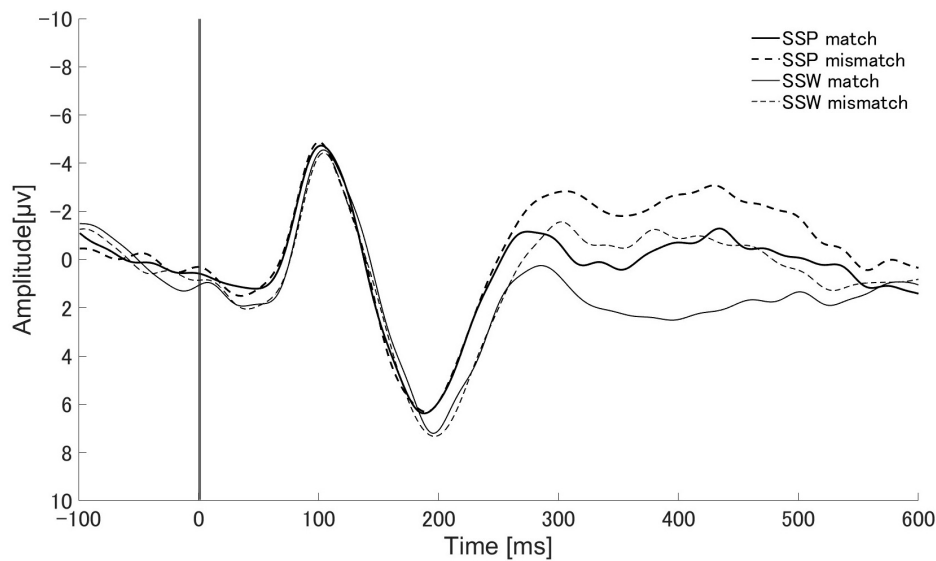


FIGURE 3

Auditory evoked responses elicited by sound symbolic words (SSW: gray lines) and sound symbolic pseudowords (SSP: black lines). Solid and dotted lines represent audio-visual matching and non-matching conditions, respectively. SSP match: match condition of sound symbolic pseudowords task. SSP mismatch: mismatch condition of sound symbolic pseudowords task. SSW match: match condition of sound symbolic words task. SSW mismatch: mismatch condition of sound symbolic words task.

There were significant differences in the amplitudes of PMN between the match and mismatch conditions for the sound symbolic words [$t(16) = 2.93, p = 0.010, \text{Cohen's } d = 0.711$] and pseudowords, [$t(16) = 2.55, p = 0.021, d = 0.619$]. Furthermore, there were significant PMN amplitude differences between the sound symbolic word and pseudoword tasks for the match condition [$t(16) = 3.07, p = 0.007, d = 0.744$] and mismatch conditions [$t(16) = 2.73, p = 0.015, d = 0.663$] (Figure 4).

Similar results were obtained for N400. We found significant N400 amplitude differences between the match and mismatch conditions for sound symbolic words [$t(16) = 6.00, p < 0.001, d = 1.455$] and pseudowords [$t(16) = 2.87, p = 0.011, d = 0.697$], and between the sound symbolic word and pseudoword tasks for the match [$t(16) = 4.66, p < 0.001, d = 1.129$] and mismatch conditions [$t(16) = 2.25, p = 0.039, d = 0.546$] (Figure 5).

There were significant differences in PMN latency between the match and mismatch conditions for sound symbolic words [$t(16) = -3.64, p = 0.002, d = -0.882$] and pseudowords [$t(16) = -2.58, p = 0.020, d = -0.626$], but no significant difference in PMN latency between the sound symbolic word and pseudoword tasks for the match [$t(16) = -0.49, p = 0.632, d = -0.119$] nor mismatch conditions [$t(16) = 0.13, p = 0.902, d = 0.030$] (Figure 6).

4. Discussion

This study investigated the semantic processing of both existing sound symbolic words and sound symbolic pseudowords with similar phonological structures using ERPs elicited while making semantic assessments of whether the sound stimuli are a match or mismatch with a picture of an event. During the task of determining whether pictures of events and speech sounds match or mismatch PMN and N400 were measured, and the semantic

processing of sound symbolic words and pseudowords was examined and compared. The results confirm the hypothesis that PMN reflects not only phonological but also word processing, while N400 reflects contextual semantic processing. Furthermore, both the sound-symbolic pseudowords (non-words) and the sound-symbolic words (existing words) showed an amplitude increase in the mismatch condition compared to the match condition, suggesting that pseudowords, like existing words, may undergo semantic processing.

Connolly et al. (2001) discovered that when phonetic features match between the target word and a candidate word, less phonological processing work is needed, and the PMN is reduced; when they do not match, more thorough phonological analysis is needed, and a larger PMN is generated. This study presented pictures presented prior to sounds, so pictures are prime for target words. The PMN results of the present study confirmed the findings of Connolly et al. (2001) by demonstrating an increase in amplitude in the mismatch condition compared to the match condition for conventional sound symbolic words. Thus, the results of this study support the hypothesis of van den Brink et al. (2001) and Kujala et al. (2004) that PMNs reflect a continuum of processing at the acoustic, phonological, and word levels, suggesting that PMNs do not represent phonological processing alone, but rather a process that reflects access to the mental lexicon.

As for the N400 following the PMN, it has been reported that when both contextual and lexical information for a word is available, the contextual information influences the N400, decreasing its amplitude in words that match the context (van Berkum et al., 1999). Therefore, Kutas and Federmeier (2000) describe the N400 as reflecting "integration into the context." Deacon et al. (2004) stated that it is generated by orthographic/phonological analysis and is affected by semantic information in a "top-down" manner. In this study, in the existing

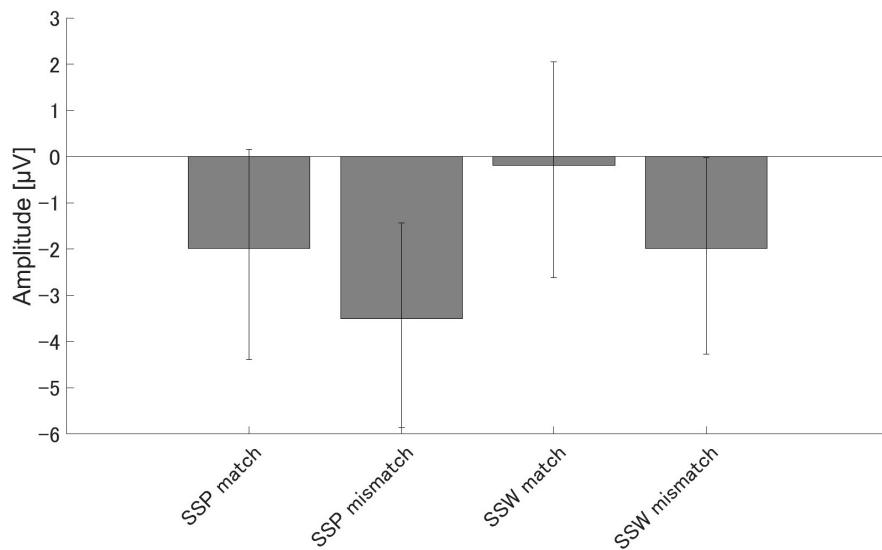


FIGURE 4

Group means ($N = 17$) of PMN amplitudes. Group means ($N = 17$) of PMN, the maximum amplitude elicited 250–350 ms after speech presentation. The error bars denote 95% confidence intervals. SSP, Sound symbolic pseudowords. SSW, sound symbolic words.

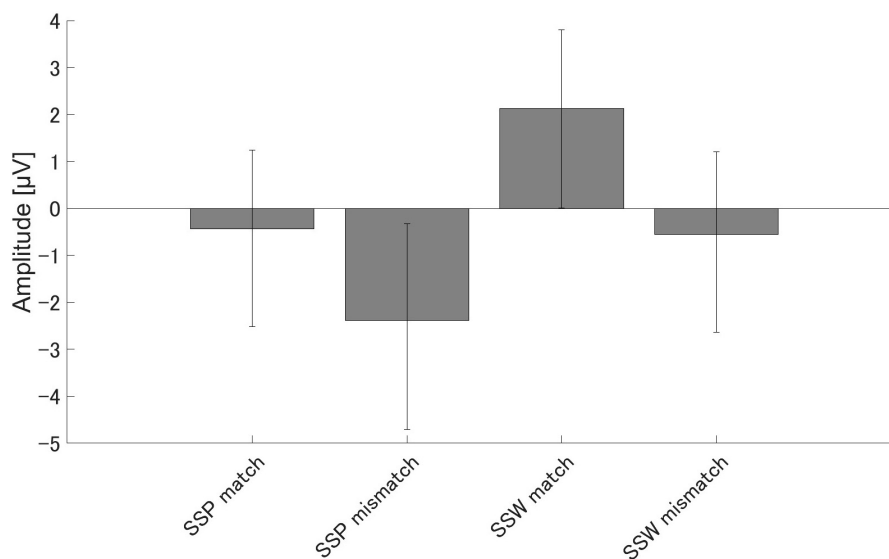


FIGURE 5

Group means ($N = 17$) of N400 amplitudes. Group means ($N = 17$) of N400 amplitudes elicited 350–500 ms after speech sound presentation. The error bars denote 95% confidence intervals.

word task, words are presented in both the match and mismatch conditions, and match/mismatch judgments must be integrated into the context of the pictures. Therefore, the difference in N400 between the match and mismatch conditions in this task may reflect not only access to the mental lexicon, but also semantic processing (integration into the context). The results obtained in this study give clues as to what processes PMN and N400 reflect: PMN may reflect the phonological and lexical processing, while N400 may reflect a higher-order processing, the process of integration into context.

As a word is being presented auditorily, several candidate words (cohort) that start with the same phoneme are activated in the

listener's mental lexicon. As the phonemes are perceived one at a time, the candidate vocabulary is reduced to the final pertinent term (Marslen-Wilson, 1987). This process has also been supported by brain wave evaluation, which discovered that for the semantic processing of vocabulary, the responses to the phoneme of the words change over time. In the McMurray et al. (2022) experiment, after target words were presented auditorily, a word phonologically similar (cohort) to the target word, and a word unrelated to the target word was presented in written form. The task was to choose which one is closer to the phonetically presented word. The results showed a characteristic pattern in which the target word and its phonologically similar competitor word (cohort) are

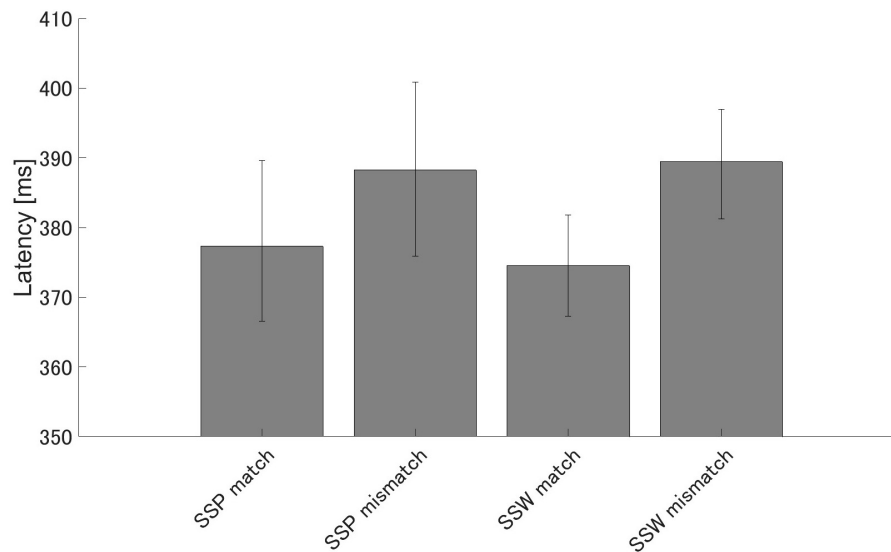


FIGURE 6

Group means ($N = 17$) of PMN latencies. Group means ($N = 17$) of PMN latencies that elicited maximum amplitude 250–350 ms after speech presentation. The error bars denote 95% confidence intervals. SSP: sound symbolic pseudowords, SSW: sound symbolic words.

active immediately after word onset (at levels that were greater than unrelated items), and at 100 ms activation levels were the same. However, 500 ms after word presentation, the response to the competing cohort is suppressed and the target is selected. If we interpret the findings of this study through the lens of the processing of words presented auditorily, PMN seems to reflect the activation of word candidates (cohorts) from the point of speech input and the process of lexical judgments in the search for the corresponding word. The subsequent N400 reflects the higher-level semantic processing of contextual integration.

Although sound symbolic pseudowords are non-words, both PMN and N400 increased in amplitude in the mismatch condition relative to the matching condition, same as in the case of existing sound symbolic words. This suggests that even though sound symbolic pseudowords are non-words, they are processed similarly to standard words in terms of semantics. Access to meaning through automatic activation of phonological information has been reported for non-words with phonological features similar to those of real words (Kujala et al., 2004). The results of this study show that this is also the case for physiological brain responses. This experiment included a situation picture and word/pseudoword matching task. The higher processing effort required for the mismatch condition compared to the match condition for the sound-symbol pseudowords suggests that word-level processing, such as retrieving word candidates, was also performed for non-word sound symbolic pseudowords. Furthermore, the increase in N400 amplitude in the mismatch condition for the sound symbolic pseudowords is consistent with Deacon et al. (2004). Therefore, it has been suggested that the sound-symbolic pseudowords are indicative of processing at a higher level, that is, semantic processing, including integration into the context. In the difference between the match and mismatch conditions for sound symbol pseudowords indicates that sound symbol pseudowords are not judged on a simple impression level (such as spiky/round) or activated by other words with similar spellings (phonological

structures) in the lexicon, but are judged on the similar context basis as actual words.

Language is the process of ascribing a label to a situation or object. It emerges as a result of semantic integration of auditory and visual information. Attentional processes are also involved (Wang et al., 2017; McCormick et al., 2018). Multiple integration processes in various cortical areas are involved in semantic integration, which is modulated by attention (Xi et al., 2019). According to McCormick et al. (2021), greater attention may be needed for the adjusting and controlling of information processing for discrepancies between visual and auditory information when attention is focused on auditory information. In this task, the concept that a word meant was presented as an illustration, and if the word matched the presented illustration, it was processed quickly and with low processing load due to the top-down effect of context. However, if the word did not match, the participants had to reinterpret the picture and search again for a possible match with the scene/event, which would have required more attention due to the processing load of adapting and suppressing information processing. As shown in Figure 1, in addition to representing the softness of the dog's fur, the picture could also represent the act of touching the dog or the feelings of the person touching the dog. It has been noted that higher-level processing of elements like syntax and contextual meaning is also carried out during the semantic processing of a word as sounds are entered one at a time and the candidate vocabulary is narrowed down; thus contextual-level semantic information has a significant impact on semantic processing (Dahan and Tanenhaus, 2004; Costa et al., 2009; McMurray et al., 2022). We assume that in the existing word task and pseudoword task, the increased amplitude and latency delay in the mismatch condition may have reflected the search for alternative interpretation of what the stimulus picture may represent, the top-down influence of different possibilities, and the load in narrowing the vocabulary. In addition, the present results showed that the amplitude of both the PMN

and the N400 was increased for sound-symbolic pseudowords (non-words) compared to sound-symbolic words (words). Sound-symbolic pseudowords are not words that are stored in the mental lexicon. Therefore, the expectation of this result was that sound-symbolic pseudowords would require more extensive processing (attention). In this study, we were able to show the differences between the semantic processing of existing sound symbolic words and pseudowords in terms of the physiological responses of the brain. The similarities between sound symbolic pseudowords and conventional words suggest that both are processed temporally from the input of phonemes to the contextual semantic processing. The difference was that the processing load for sound symbolic pseudowords was greater than that for sound symbolic words. In order to clarify this difference, it will be necessary to identify the brain regions involved in the semantic processing of both sound symbolic words and pseudowords, which is an issue for the future. Although fMRI at high spatial resolution or other methods could be used to search for the region of the brain responsible for semantic processing of sound symbolic pseudowords, it is challenging to distinguish PMN from N400 due to inferior temporal resolution. Therefore, a more detailed analysis using a combination of various neuroimaging techniques is needed. In addition, Japanese is a language that is rich in onomatopoeia and ideophones and commonly makes use of sound symbolism. This could be why sound symbolic words are processed in a way similar to the semantic processing of other words of the language. Thus, a cross-linguistic approach is necessary for future studies in order to clarify the semantic processing of sound symbolism.

5. Conclusion

In this study, the semantic processing of sound symbolic words included in the mental lexicon is compared to the semantic processing of sound symbolic pseudowords, which have no concrete meaning stored in the lexicon, but use sounds that are consistently associated with a certain meaning. It also clarifies the processes reflected by PMN and N400, which are used in many studies of lexical semantic processing. The findings of this study suggest that PMNs may reflect not only phonological processing, but also further processing up to the point of accessing vocabulary, whereas N400 may reflect final semantic judgments. Furthermore, similar responses toward sound symbolic words and pseudowords were observed in the match and mismatch conditions, indicating that both of them undergo semantic processing.

Data availability statement

The original contributions presented in this study are included in the article/**Supplementary material**, further inquiries can be directed to the corresponding author.

Ethics statement

The studies involving humans were approved by the International University of Health and Welfare's Ethical Review Committee. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

KS and HO contributed to conception and design of the study. KS and SK performed the experiments, collected the data, and performed the statistical analysis. KS wrote the first draft of the manuscript. KS, SK, JI, MP and HO wrote sections of the manuscript. MP edited English language. All authors contributed to manuscript revision, read, and approved the submitted version.

Funding

This work was supported by the Japan Society for the Promotion of Science (JSPS) Grants in Aids for Scientific Research (KAKENHI) Grant Number 21K13620.

Acknowledgments

We gratefully acknowledge our participants for their diligent cooperation and the work of present members of our laboratory.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fnhum.2023.1208572/full#supplementary-material>

References

- Abe, J., Momouti, Y., and Kaneko, Y. (1994). *Human language information processing - Cognitive Science of Language Understanding*. Tokyo: Science Inc.
- Aitchison, J. (2012). *Words in the Mind: An Introduction to the Mental Lexicon*, 4th Edn. New Jersey: Wiley-Blackwell.
- Akita, K. (2017). "Do foreign languages have onomatopoeia?" in *The Mystery of Onomatopoeia - From Pikachu to Mohammed - Iwanami Science Library 261*, ed. H. Kubozono (Tokyo: Iwanamisyoten).
- Asano, M., Imai, M., Kita, S., Kitajo, K., Okada, H., and Thierry, G. (2015). Sound symbolism scaffolds language development in preverbal infants. *Cortex* 63, 196–205. doi: 10.1016/j.cortex.2014.08.025
- Bakeman, R. (2005). Recommended effect size statistics for repeated measures designs. *Behav. Res. Methods* 37, 379–384. doi: 10.3758/BF03192707
- Brown, C., and Hagoort, P. (1993). The processing nature of the n400: evidence from masked priming. *J. Cogn. Neurosci.* 5, 34–44. doi: 10.1162/jocn.1993.5.1.34
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. New York, NY: Routledge Academic.
- Connolly, J. F., Service, E., D'Arcy, R. C., Kujala, A., and Alho, K. (2001). Phonological aspects of word recognition as revealed by high-resolution spatio-temporal brain mapping. *Neuroreport* 12, 237–243. doi: 10.1097/00001756-200102120-00012
- Costa, A., Strijkers, K., Martin, C., and Thierry, G. (2009). The time course of word retrieval revealed by event-related brain potentials during overt speech. *Proc. Natl. Acad. Sci. U. S. A.* 106, 21442–21446. doi: 10.1073/pnas.0908921106
- Dahan, D., and Tanenhaus, M. K. (2004). Continuous mapping from sound to meaning in spoken-language comprehension: immediate effects of verb-based thematic constraints. *J. Exp. Psychol.* 30, 498–513. doi: 10.1037/0278-7393.30.2.498
- D'Arcy, R. C., Connolly, J. F., Service, E., Hawco, C. S., and Houlihan, M. E. (2004). Separating phonological and semantic processing in auditory sentence processing: a high-resolution event-related brain potential study. *Hum. Brain Mapp.* 22, 40–51. doi: 10.1002/hbm.20008
- Deacon, D., Dynowska, A., Ritter, W., and Grose-Fifer, J. (2004). Repetition and semantic priming of nonwords: Implications for theories of N400 and word recognition. *Psychophysiology* 41, 60–74. doi: 10.1111/1469-8986.00120
- Desroches, A. S., Newman, R. L., and Joanisse, M. F. (2009). Investigating the time course of spoken word recognition: electrophysiological evidence for the influences of phonological similarity. *J. Cogn. Neurosci.* 21, 1893–1906. doi: 10.1162/jocn.2008.21142
- Dumay, M., Benraïss, A., Barriol, B., Colin, C., Radeau, M., and Besson, M. (2001). Behavioral and electrophysiological study of phonological priming between bisyllabic spoken words. *J. Cogn. Neurosci.* 13, 121–143. doi: 10.1162/089892901564117
- Hagoort, P., Hald, L., Bastiaansen, M., and Petersson, K. M. (2004). Integration of word meaning and world knowledge in language comprehension. *Science* 304, 438–441. doi: 10.1126/science.1095455
- Hamano, S. (2014). *Onomatopoeia in Japanese: Sound Symbolism and Structure*. Tokyo: Kurosio Publishers.
- Hinton, L., Nichols, J., and Ohala, J. (1994). *Sound symbolism*. Cambridge: Cambridge University Press.
- Hockett, C. F. (1960). The origin of speech. *Sci. Am.* 203, 89–96.
- Köhler, W. (1947). *Gestalt psychology*, 2nd Edn. New York: Liveright Publishing.
- Kovic, V., Plunkett, K., and Westermann, G. (2010). The shape of words in the brain. *Cognition* 114, 19–28. doi: 10.1016/j.cognition.2009.08.016
- Kujala, A., Alho, K., Service, E., Ilmoniemi, R. J., and Connolly, J. F. (2004). Activation in the anterior left auditory cortex associated with phonological analysis of speech input: Localization of the phonological mismatch negativity response with MEG. *Brain Res. Cogn. Brain Res.* 21, 106–113. doi: 10.1016/j.cogbrainres.2004.05.011
- Kutas, M., and Federmeier, K. D. (2000). Electrophysiology reveals semantic memory use in language comprehension. *Trends Cogn. Sci.* 4, 463–470. doi: 10.1016/s1364-6613(00)01560-6
- Kutas, M., and Federmeier, K. D. (2011). Thirty years and counting: Finding meaning in the N400 component of the event-related brain potential (ERP). *Annu. Rev. Psychol.* 62, 621–647. doi: 10.1146/annurev.psych.093008.131123
- Kutas, M., and Hillyard, S. A. (1980). Reading senseless sentences: brain potentials reflect semantic incongruity. *Science* 207, 203–205. doi: 10.1126/science.7350657
- Lee, J. Y., Harkrider, A. W., and Hedrick, M. S. (2012). Electrophysiological and behavioral measures of phonological processing of auditory nonsense V-CV-VCV stimuli. *Neuropsychologia* 50, 666–673. doi: 10.1016/j.neuropsychologia.2012.01.003
- Manfredi, M., Cohn, N., and Kutas, M. (2017). When a hit sounds like a kiss: An electrophysiological exploration of semantic processing in visual narrative. *Brain Lang.* 169, 28–38. doi: 10.1016/j.bandl.2017.02.001
- Marslen-Wilson, W. D. (1987). Functional parallelism in spoken word-recognition. *Cognition* 25, 71–102. doi: 10.1016/0010-0277(87)90005-9
- McCormick, K., Lacey, S., Stilla, R., Nygaard, L. C., and Sathian, K. (2018). Neural basis of the crossmodal correspondence between auditory pitch and visuospatial elevation. *Neuropsychologia* 112, 19–30. doi: 10.1016/j.neuropsychologia.2018.02.029
- McCormick, K., Lacey, S., Stilla, R., Nygaard, L. C., and Sathian, K. (2021). Neural basis of the sound-symbolic crossmodal correspondence between auditory pseudowords and visual shapes. *Multisens. Res.* 35, 29–78. doi: 10.1163/22134808-bja10060
- McMurray, B., Sarrett, M. E., Chiu, S., Black, A. K., Wang, A., Canale, R., et al. (2022). Decoding the temporal dynamics of spoken word and nonword processing from EEG. *Neuroimage* 260:119457. doi: 10.1016/j.neuroimage.2022.119457
- Newman, R. L., and Connolly, J. F. (2009). Electrophysiological markers of pre-lexical speech processing: evidence for bottom-up and top-down effects on spoken word processing. *Biol. Psychol.* 80, 114–121. doi: 10.1016/j.biopsycho.2008.04.008
- Newman, R. L., Connolly, J. F., Service, E., and McIvor, K. (2003). Influence of phonological expectations during a phoneme deletion task: evidence from event-related brain potentials. *Psychophysiology* 40, 640–647. doi: 10.1111/1469-8986.00065
- Ono, M. (2007). *Imitative words and Mimetic words 4500 Japanese Onomatopoeia Dictionary*. Tokyo: SHOGAKUKAN.
- Osaka, M. (1999). "Multilingual comparison of onomatopoeia and ideophone," in *Exploration of the language of sensibility - Where is the heart in onomatopoeia and ideophone?*, ed. N. Osaka (Tokyo: Shinyousya).
- Petten, C. V., and Kutas, M. (1990). Interactions between sentence context and word frequency in event-related brain potentials. *Mem. Cogn.* 18, 380–393. doi: 10.3758/bf03197127
- Pinker, S. (1999). *Words and Rules: the ingredients of language*. New York: Basic Books.
- Ramachandran, V., and Hubbard, E. (2001). Synaesthesia – A window into perception, thought and language. *J. Conscious. Stud.* 8, 3–34.
- Sidhu, D. M., and Pexman, P. M. (2018). Five mechanisms of sound symbolic association. *Psychon. Bull. Rev.* 25, 1619–1643.
- Stille, C. M., Bekolay, T., Blouw, P., and Kröger, B. J. (2020). Modeling the mental lexicon as part of long-term and working memory and simulating lexical access in a naming task including semantic and phonological cues. *Front. Psychol.* 11:1594. doi: 10.3389/fpsyg.2020.01594
- Sučević, J., Savić, A. M., Popović, M. B., Styles, S., and Ković, V. (2015). Balloons and bavoons versus spikes and shikes: ERPs reveal shared neural processes for shape-sound-meaning congruence in words, and shape-sound congruence in pseudowords. *Brain Lang.* 145, 11–22. doi: 10.1016/j.bandl.2015.03.011
- Tamori, I., and Schourup, L. (1999). *Onomatopoeia - form and meaning - Japanese-English Subject Matter Study Series*. Tokyo: Kuroshio Publishing.
- van Berkum, J. J., Hagoort, P., and Brown, C. (1999). Semantic integration in sentences and discourse: Evidence from the N400. *J. Cogn. Neurosci.* 11, 657–671. doi: 10.1162/089892999563724
- van den Brink, D., Brown, C. M., and Hagoort, P. (2001). Electrophysiological evidence for early contextual influences during spoken-word recognition: N200 versus N400 effects. *J. Cogn. Neurosci.* 13, 967–985. doi: 10.1162/089892901753165872
- Wang, H., Zhang, G., and Liu, B. (2017). Influence of auditory spatial attention on cross-modal semantic priming effect: Evidence from N400 effect. *Exp. Brain Res.* 235, 331–339. doi: 10.1007/s00221-016-4792-4
- Westbury, C. (2005). Implicit sound symbolism in lexical access: Evidence from an interference task. *Brain Lang.* 93, 10–19. doi: 10.1016/j.bandl.2004.07.006
- Xi, Y., Li, Q., Gao, N., He, S., and Tang, X. (2019). Cortical network underlying audiovisual semantic integration and modulation of attention: An fMRI and graph-based study. *PLoS One* 14:e0221185. doi: 10.1371/journal.pone.0221185