

Neural basis of morphologically complex word processing: EEG studies

中島, 潤

<https://hdl.handle.net/2324/6787384>

出版情報 : Kyushu University, 2022, 博士 (文学) , 課程博士
バージョン :
権利関係 :



**Neural basis of morphologically
complex word processing:
EEG studies**

中島 潤

Jun NAKAJIMA

A Dissertation submitted to the Department of Linguistics
in Partial Fulfilment of the Requirements
for the Doctoral Degree in Linguistics

Kyushu University

Table of Contents

List of Tables	v
List of Figures	vi
Abstract	1
1. Introduction	3
2. Neurophysiological methods and brain activations	12
2.1 Methods	13
2.1.1 Electroencephalography (EEG)	13
2.1.2 Magnetoencephalography (MEG)	15
2.2 Brain activations	16
2.2.1 The N170/M170: Decomposition	16
2.2.2 The N400/M350: Lexical processing	19
3. Experiment 1: Lexical decision in the <i>hiragana</i> script..	21
3.1 Background	22
3.1.1 Introduction	22
3.1.2 Predictions	26
3.2 Experiment 1: Lexical decision in <i>hiragana</i> script	26
3.2.1 Participants	26
3.2.2 Stimuli	27
3.2.3 Procedure	30
3.2.4 EEG data acquisition	31
3.2.5 Data analysis	32
3.2.6 Data processing	32
3.3 Results	35
3.3.1 Behavioral results	35
3.3.2 ERP results	37
3.3.2.1 N170: Morphological decomposition	37
3.3.2.2 N400: Lexical processing	40
3.4 Discussion	41

3.5 Conclusion.....	51
4. Experiment 2: Masked priming.....	53
4.1 Background	54
4.1.1 Introduction	54
4.1.2 Experimental paradigm: Masked priming	55
4.1.3 Predictions	58
4.2 Experiment 2: Masked priming.....	60
4.2.1 Participants	60
4.2.2 Stimuli	61
4.2.3 Procedure	65
4.2.4 EEG data acquisition	66
4.2.5 Data analysis	66
4.2.6 Data processing	67
4.3 Results	73
4.3.1 Results 1: <i>-mi</i> type (<i>-sa</i> and <i>-mi</i> nouns)	74
4.3.1.1 Behavioral results 1: <i>-mi</i> type (<i>-sa</i> and <i>-mi</i> nouns)	75
4.3.1.2 ERP results 1: <i>-mi</i> type (<i>-sa</i> and <i>-mi</i> nouns)	78
4.3.1.2.1 Priming effects on the N400	78
4.3.1.2.2 Attenuation of the N400 by TP.....	84
4.3.1.2.3 Significant laterality effects on the N170	86
4.3.2 ERP Results 2: non- <i>mi</i> type (<i>-sa</i> nouns)	90
4.3.2.1 Priming effects on the N400	90
4.3.2.2 Attenuation of the N400 by TP	94
4.4 Discussions.....	99
4.4.1 <i>-mi</i> type (<i>-sa</i> and <i>-mi</i> nouns)	99
4.4.2 non- <i>mi</i> type (<i>-sa</i> nouns)	104
4.5 Conclusions	109
4.5.1 <i>-mi</i> type (<i>-sa</i> and <i>-mi</i> nouns)	109
4.5.2 non- <i>mi</i> type (<i>-sa</i> nouns)	111
5. Experiment 3: Cortical Entrainment.....	113
5.1 Background	114
5.1.1 Introduction	118
5.2 Experiment	120
5.2.1 Participants	120

5.2.2	Stimuli	121
5.2.3	Procedure	123
5.2.4	EEG data acquisition	126
5.2.5	Data analysis	126
5.2.5.1	Behavioral data analysis	127
5.2.5.2	EEG data analysis	127
5.2.6	Data processing	128
5.3	Results	129
5.3.1	Behavioral results	129
5.3.2	EEG results	131
5.3.3	Auditory results	131
5.3.4	Visual results	137
5.3.5	Comparison between auditory and visual modalities	142
5.4	Discussion	143
5.5	Conclusion.....	148
6.	General Discussion	150
6.1	Morphological decomposition of complex words	151
6.2	Lexical processing of complex words.....	152
6.3	Cortical entrainment.....	154
7.	Conclusions	156
	References	163
	Acknowledgements.....	167
	Appendices	172

List of Tables

Table 1. Summary statistics of experimental stimuli (experiment 1) .	30
Table 2. ANOVA results for accuracy (experiment 1)	36
Table 3. ANOVA results for reaction times (experiment 1).....	37
Table 4. The comparison among LME models	39
Table 5. A summary of neurolinguistic studies demonstrating the correlation between N/M170 and TP of affixes.....	45
Table 6. Summary statistics of experimental stimuli (experiment 2) .	64
Table 7. Behavioral data and priming effects (- <i>mi</i> type).....	74
Table 8. ANOVA results for accuracy	75
Table 9. ANOVA results reaction time (- <i>mi</i> type).....	77
Table 10. ANOVA results of the N400 (- <i>mi</i> type).....	81
Table 11. Statistical results of averaged N170 amplitude for - <i>mi</i> type	89
Table 12. ANOVA results of the N400 (non- <i>mi</i> type).....	93
Table 13. Summary of similarities and dissimilarities between - <i>sa</i> and - <i>mi</i> nouns	100
Table 14. Prediction for the evoked power of the frequency of interest	118

List of Figures

Figure 1. Experimental procedure.....	31
Figure 2. aROIs of the N170	37
Figure 3. Grand averages of the N170 among the four factors.....	38
Figure 4. Beta (estimate) of TP for the Best LME model	39
Figure 5. aROIs of the N400	40
Figure 6. Grand averages of the N400 among the four factors.....	41
Figure 7. Experimental procedure using the masked priming paradigm	65
Figure 8. ROIs for N400 (A) and N170 (B).....	70
Figure 9. Grand-averaged ERPs of the four factors of the <i>-mi</i> type ...	80
Figure 10. The scalp distribution of the unrelated – related	80
Figure 11. Mean amplitude of N400 between conditions for each ROIs (Prime type: black and white, Suffix: right and left)	83
Figure 12. Effect of the TP between Stem and Suffix in the N400 regions	85
Figure 13. Significant laterality effects on the N170 in the temporal and parieto-occipital regions.....	88
Figure 14. Grand-averaged ERPs of the four factors of <i>-mi</i> type	92
Figure 15. Effect of the No. of morpheme for non- <i>mi</i> type in the N400 regions	95
Figure 16. Effect of the TP for non- <i>mi</i> type in the N400 regions	98
Figure 17. Schematic structures of word stimuli	117
Figure 18. Normalized evoked power (grand average over all channels) of auditory stimuli for four conditions (log scale)....	134
Figure 19. The spectral activity across four conditions for normalized evoked power in the auditory experiment.....	135
Figure 20. Normalized evoked power (grand average over all channels) of visual stimuli for four conditions (log scale).....	139
Figure 21. The spectral activity across four conditions for normalized evoked power in the visual experiment.....	140

Abstract

Language is a unique cognitive faculty of humans. This faculty allows us to generate an infinite range of expressions (i.e., phrases and sentences) from a finite set of smaller units that bear meanings (i.e., words). Therefore, one of the core problems in linguistics is the mechanisms between phrases/sentences (syntax) and words (morphology). Unlike phrases/sentences, which are not limited in their productivity, words are limited in their productivity and irregularity (e.g., allomorphs). Due to these features, words have been dealt with differently in phrases/ sentences.

Previous theoretical and neurolinguistic studies have investigated processing phrases and sentences, revealing their processing profile in the brain, and theoretical studies of words suggested that words have a similar hierarchical structure of phrases/sentences, suggesting that a similar syntactic operation can explain both syntax and morphology. However, this fundamental implication of linguistics has not been well investigated in the neurophysiological aspect, and research on word processing in Japanese has inconsistent results.

In this dissertation, we fill the following two gaps in word processing with three EEG experiments: the inconsistent results of the processing of de-adjectival nouns in Japanese and the similarity in processing the words

and phrases/sentences with their hierarchical structure. This dissertation is organized as follows.

Chapter 1 starts the dissertation by reviewing the research history of the processing of complex words. An overview of the dissertation is also provided. In Chapter 2, we introduce experimental methods and terms of brain activation to interpret the dissertation. In Chapters 3 and 4, we solve the contradiction of de-adjectival nouns in Japanese through two EEG experiments. Specifically, focusing on functionally different stages of word processing [i.e., morphological decomposition (Chapter 3) and lexical processing (Chapter 4)], we demonstrate that the de-adjectival nouns are treated as morphemes but not whole words. Then, we fill the gap between word processing and phrase/sentence processing by demonstrating that the hierarchical structure of words evokes a similar brain activation of phrases and sentences (Chapter 5). In Chapter 6, we discuss the results for their implications. First, we demonstrate the cross-linguistic universality of word processing from the results of Chapters 3 and 4. Then, we demonstrate a common neural mechanism between words and phrases/sentences from the results of Chapter 5. Finally, in Chapter 7, we summarize the results and refer to limitations and propose future directions in the processing of words.

1. Introduction

In theoretical linguistics, phrases and sentences are linguistic units represented as hierarchical structures generated by syntactic rules (Chomsky, 1957). The smaller linguistic unit, words, have been considered fundamentally different from phrases/sentences as words have unproductive/irregular derived or inflected forms, unlike phrases and sentences (e.g., curious, *cury, curiosity, & curiousness, vs. furious, fury, *furiosity, & furiousness; Aronoff, 1976). From this view, similar generative syntax rules cannot explain these differences in productivity/irregularity since they have unexpectable grammatical forms (e.g., allomorph), as mentioned above. For this reason, the morphology that governs the words is often treated separately from syntax.

To explain this phenomenon, three theories have been provided to explain how words and their components' morphemes are processed by neural mechanisms. The first theory is the single mechanism or storage (full listing) theory, which assumes that all words are memorized in the whole word form in the mental lexicon (Butterworth, 1983). This theory does not assume any morphemes. The second theory is the dual mechanism theory (Pinker & Prince, 1988; Pinker 1991), which assumes two processing routes for word processing: productive words are generated by rule (e.g., attaching a morpheme "ed" to a verb; e.g., walk → walk-ed), whereas

irregular/unproductive words (e. g., allomorphs) are memorized as whole word form in the mental lexicon (e. g., fall → fell, *fall-ed). This theory assumes morphemes for productive words but not for irregular/unproductive words. The third theory is the single mechanism composition theory, which assumes that all words are decomposed into abstract morphemes (e. g., fall → fell = (fall [+PAST])) regardless of their irregularity/productivity of words (Taft & Forster, 1975). Halle and Marantz (1993) propose a more detailed theory called distributed morphology. The theory assumes that hierarchical structures consist of abstract morphemes for all mono/polymorphemic words regardless of their irregularity/productivity. Furthermore, the theory assumes a similar syntactic operation for both words and phrases/sentences to make their structures (the single-engine hypothesis). In other words, the theory integrates morphology and syntax to explain a similar cognitive and neural mechanism for words and phrases/sentences.

To examine which theory is plausible, many behavioral and neurolinguistic studies have been conducted (Leminen et al., 2018). These studies mainly dealt with the irregular verb allomorph of English past tense (e.g., walk & walk-ed vs. fall & fell), leading to the “past tense debate” (Pinker & Ullman, 2002; Fruchter et al., 2013). In a behavioral study, Marslen-Wilson (1993) provided negative evidence of the single mechanism

storage theory using cross-modal priming. Besides, Marslen-Wilson and Tyler (1998) replicated the result using long lag priming, supporting the dual mechanism theory or single mechanism composition theory. In addition, the study demonstrated similar results between regular and irregular past tense, supporting the single mechanism composition theory. Moreover, the “wug test” conducted by Albright and Hayes (2003), which dealt with the production of novel irregular verbs, also demonstrated the plausibility of the third theory.

In a neurolinguistic study, Stockall and Marantz (2006) reported that lexical access (retrieval of meanings) to the verbal root occurs in both regular and irregular verbs. Furthermore, Fruchter et al. (2013) demonstrated that the irregular verb is recognized as two morphemes. These studies elucidated that irregular verbs are processed/recognized as a composition of abstract verbal root and past tense morpheme “ed” with non-phonetic features and provided further evidence for the single mechanism composition theory.

In summary, these behavioral and neurolinguistic studies elucidated that the single mechanism composition theory can be used to process morphologically complex words¹ Moreover, a related study for derivation replicated similar results (Solomyak & Marantz, 2009; Gwilliams et al.,

¹ Henceforth, the term “(morphologically) complex” refers to that inflected, derived, or compound words regardless of their regularity/productivity.

2016). These studies of irregular/unproductive inflections and derivations suggested that, as well as phrases and sentences, words have a common neural mechanism that syntactic operation can explain, such as allomorphs. Further, these results imply that the words have hierarchical structures, such as phrases/sentences. Moreover, Halle and Marantz (1993) proposed the single-engine hypothesis: there is only one generative engine to build complex words and sentences. That is, phrases/sentences and words have common hierarchical structures built by a fundamental syntactic operation, Merge.

Nevertheless, this fundamental implication of linguistics, a similar operation between words and phrases/sentences, is yet to be examined experimentally. Besides, studies in two types of Japanese de-adjectival nouns offer inconsistent results between the two theory of word processing. The two types of Japanese de-adjectival nouns differ in their lexicalized meanings and productivities. In the study of these nouns, Hagiwara et al. (1999) supported the dual mechanism theory, whereas Clahsen and Ikemoto (2012) and Fiorentino et al. (2016) supported the single mechanism composition theory. However, this inconsistency is difficult to interpret due to the different experimental paradigms among these studies.

This dissertation fills the two gaps in morphologically complex word

processing through neurophysiological experiments. The aim of this dissertation is as follows:

- (1) To solve the inconsistent results of word processing in Japanese de-adjectival nouns (Chapters 3 & 4: the single mechanism composition theory).
- (2) To demonstrate a common neural mechanism between words and phrases/sentences. (Chapter 5: the single-engine hypothesis).

This dissertation is organized as follows.

Chapter 2 introduces the methods for neurophysiological study and essential terms for the dissertation.

Chapter 3 presents an electroencephalography (EEG) experiment to solve the inconsistency of de-adjectival nouns in Japanese by a lexical decision task. Specifically, given the previous literature, we examine whether the transition probability between morphemes predicts the evoked component called N170, which reflects the decomposition of words into morphemes. The results demonstrate that the transition probability correlates with the amplitude of N170, suggesting that both de-adjectival nouns are decomposed into morphemes, not whole word (3.3.2.1.). The results support the single mechanism composition theory and imply that the morphological processing of words has language universal neural mechanisms. Besides, the

results suggest that the unique mixed Japanese writing systems (i.e., *hiragana* and *kanji* scripts) assist the latter processing stage, the lexical processing of words (3.3.2.2.).

Chapter 4 reports an additional EEG study examining the other stage of word processing, lexical processing (retrieval and composition of meanings). Specifically, using a masked priming paradigm, we examined whether transition probability also predicts the evoked response component of the N400, which reflects lexical processing. The results of the experiment demonstrate that, as well as the morphological decomposition (N170), the transition probability correlates with the amplitude of N400, suggesting that both de-adjectival nouns are also treated as morphemes in the stage of the lexical processing (4.3.1.). The results suggest that the de-adjectival nouns are not treated as whole words, replicating the results of Chapter 3 and previous psycholinguistic studies. In addition, we found the difference between the two de-adjectival nouns in lexicalized meanings for the N400, demonstrating that the inconsistency between the two nouns is caused by the experimental paradigms. Moreover, further analyses demonstrate that lexical processing has two stages, lexical access and morphological composition. Furthermore, additional analyses of Japanese de-adjectival nouns reveal that the transition probability and morphological complexity affect the two stages

of lexical processing (i.e., lexical access and morphological composition) in opposite ways (4.3.2.).

In light of previous chapters, Chapter 5 demonstrates a common neural mechanism between words and phrases/sentences. Specifically, we demonstrate that the isochronous rhythmic auditory presentation of complex words with linear structure induces cortical (brain/neural)² activation that synchronizes the words' isochronous presentation of hierarchical structures (5.3.3.). The result replicates the previous studies' isochronous auditory presentation of phrases and sentences. The result strongly suggests that a common syntactic operation processes the word and phrases/sentences in the brain, which supports distributed morphology through an experimental aspect. Moreover, the cortical synchrony is also observed in the visual presentation (5.3.4.). The two results demonstrate that the neural mechanisms of word processing are modality-independent, implying that, regardless of words and phrases/sentences, auditory or visual input of linguistic structures is processed through a common neural mechanism.

In Chapter 6, we discuss the results for their implications. First, we demonstrate the cross-linguistic universality of word processing from the results of Chapters 3 and 4 (the single mechanism composition theory). Then,

² In the dissertation, we do not distinguish the terms neural, brain, or cortical (cortex) activation due to methodological limitations. See section 2.1.1.

we demonstrate a common neural mechanism between words and phrases/sentences from the results of Chapter 5, suggesting the validity of the single-engine hypothesis.

Finally, in Chapter 7, we summarize the results, refer to limitations, and propose future directions in processing words.

2. Neurophysiological methods and brain activations

In this chapter, we introduce neuroimaging methods and important brain (cortical/neural) activations widely used to elucidate the visual processing of complex words. We review their features and how they can be interpreted.

2.1 Methods

2.1.1 Electroencephalography (EEG)

Electroencephalography (EEG) is a neuroimaging method that records changes in electrical neural activities taken from the scalp by electrodes. The electrodes of the EEG are located parallel to the gyri (i.e., perpendicularly to the sulci), and they record electric changes of gyri through the surrounding tissues (i.e., spinal liquid, meninges, skull, and scalp). However, the EEG has relatively poor temporal resolution compared to the other neuroimaging methods because of the distortion and attenuation of the electric change by these heterogeneous tissues. Due to this problem, it is impossible to determine the accurate location of a specific neural activity, so the spatial resolution of the EEG is limited compared to other neuroimaging techniques, such as functional magnetic resonance imaging (fMRI), positron emission tomography (PET), and magnetoencephalography (MEG). Electrocorticography (ECoG) is a type of EEG that places electrodes on the

cerebral cortex. As there is no distortion or attenuation of the EEG, the ECoG offers higher spatial resolution (approximately 1 cm) instead of its high invasiveness.

Event-related potentials (ERPs) are neural activities induced by a specific cognitive event or stimulus, such as the presentation of words recorded by the EEG. Due to their small potentials (approximately 10 μV) and considerable noises (e.g., eye blinks and, cardiac beat line noise), the ERPs are calculated by averaging many trials to reduce the effects of these noises. ERPs are distinguished by following several features, polarity (negative or positive), peak latency (the peak from the stimulus), amplitude (voltage change: μV), and scalp distribution, and named from two features: their polarity (“N” for negative and “P” for positive) and the peak latency for a three-digit number (e.g., N170, N400).

EEG bands (waves) are rhythmic oscillatory activities divided by frequency (i.e., rhythmical periodic cycles per second). The primary waves are the Delta band (0.5–4 Hz), Theta band (4–7 Hz), Alpha band (8–12 Hz), Beta band (13–30 Hz), low Gamma band (30–50 Hz), and high gamma band (> 50 Hz). These rhythmic oscillatory activities at different frequencies have been associated with distinct stages of cognitive processing and the connection of brain areas (Buzsáki & Draguhn, 2004). In the neurolinguistic

study, the different frequency bands of the cortical entrainment have different subfunctions of cortical information processing (e.g., processing of phonemes, syllables, phrases, and sentences) (Meyer, 2018; Martin & Dumas, 2017; Martin, 2020).

2.1.2 Magnetoencephalography (MEG)

Magnetoencephalography (MEG) is another neuroimaging technique that records changes in a magnetic field produced by electrical activity. The sensors of the MEG are located parallel to the gyri (i.e., perpendicularly to the sulci). The magnetic field change occurs perpendicularly to the gyri (i.e., parallel to the gyri). Therefore, in principle, MEG measures neural activation in the sulci but not the gyri in the brain. Unlike EEG, MEG has high spatial resolution because the recorded changes in a magnetic field are not distorted by tissue. Contrasting to the other technique that records brain activity (e.g., fMRI: functional magnetic resonance imaging, NIRS: near-infrared spectroscopy), EEG and MEG have a high temporal resolution, so they are often compared to elucidate the temporal profile of a cognitive role associated with brain activation. MEG components are named with an “M” and a three-digit number indicating their time after the onset of the stimulus (e.g., M170, M350).

2.2 Brain activations

In this section, we introduce two ERP components, N170 and N400, related to visual processing (morphologically complex) words to review how they relate to processing complex words. Previous EEG and MEG studies revealed that the processing of complex words has two main stages: decomposition and lexical processing in this order. The two ERPs and their corresponding MEG components demonstrate how complex words are processed. The N170 (M170 for the MEG component) reflects decomposition. The other component, the N400 (M350 for the MEG component), reflects lexical processing. The word length and the log lemma transition probability (logarithmic co-occurrence probability between two morphemes) are important variables as it correlates with the activation of the two EEG/MEG components.

2.2.1 The N170/M170: Decomposition

The N170³, and its corresponding component of MEG M170, is the neural activation peaking at around 170 ms after the visual word onset in the

³ The N170/M170 component is sometimes mentioned as the N250/M250 in several MEG and EEG studies because it is sometime measured from the onset time of the pre-stimulus events (e.g., Morris et al., 2007 and its subsequent studies Morris & Stockall, 2012; Leminen et al., 2019; Wray et al., 2022).

left fusiform and inferior temporal gyri (L. FG/ITG). The component reflects automatic morpho-orthographical decomposition (i.e., morpheme recognition based on the exhaustive parsability⁴ of the word) and is sensitive to the length of the word.

In an MEG study of derived words, Zweig and Pytkänen (2009) demonstrated that derived bimorphemic words (e.g., “farm-er”⁵) evoked larger activation than monomorphemic words (e.g., “switch”) in the left FG/ITG using MEG. In addition, another MEG study revealed that log lemma transition probabilities (henceforth TP) between morphemes were correlated with the amplitude of M170 but not whole word frequency (Solomyak & Marantz, 2010). These results elucidated that the N170 is a reliable indicator for morphological decomposition. Moreover, as well as genuine complex words, a subsequent study demonstrated that pseudo-complex words (e.g., the word “brother” can be parsed into existing morphemes “broth” and “er”) had a similar correlation in a subsequent MEG study (Lewis et al., 2011). The result demonstrated that the N170 decompositional process is garden path-like based on exhaustive parsability.

⁴ The term “exhaustive parsability” indicates whether a word is parsed into existing morphemes. For instance, as mentioned in the main text, “teacher” and “brother” are exhaustive parsable since the morphemes “teach,” “broth,” and “er” are existing morphemes. However, “brothel” is not exhaustive parsable since “el” is not an existing morpheme. See more details for Lewis, Solomyak & Marantz, (2011) and Gwilliams & Marantz (2018).

⁵ In this dissertation, we use hyphens (“-”) for a (possible) morpheme boundary (e.g., walk-ed and farm-er).

In an EEG study, Lavric et al. (2007) reported similar differences between bimorphemic and monomorphemic words in the time window of N170 in an ERP study. These studies revealed that the N170/M170 reflects the processing cost of morpho-orthographical decomposition. The TP is a reliable index for whether the neural system of visual word processing considers a word morphologically complex.

In a MEG study of English inflection, Fruchter et al. (2013) found that the M170 evoked larger activation for the pairs of the past and the present tense of irregular verbs (e.g., fell & fall) as well as the past tense of regular verbs (e.g., walked & walk), while orthographical controls did not (e.g., bell & ball). The study demonstrated that the past tense form is recognized as two abstract morphemes (i.e., stem and past tense morpheme “ed”) even though the verb form cannot be decomposed orthographically or phonologically. A similar result is reported in the EEG (Morris & Stockall, 2012). Other neurolinguistic studies reported similar correlations between TP and activations cross-linguistically [Greek: Neophytou et al., 2018 (derivation); Finnish: Hakala et al., 2018 (derivation and inflection); Japanese: Ohta et al., 2019 (inflection); Tagalog: Wray et al., 2022 ⁶ (derivation)]. These neurolinguistic studies revealed that the N170/M170 reflects morpho-

⁶ The study also deals with infixation, circumfixation, and reduplication in Tagalog.

orthographical and abstract-level decomposition in the visual modality.

2.2.2 The N400/M350: Lexical processing

Another neural activation, called N400/M350, is evoked around 300–500 ms after the visual word onset, which peaks around 350 ms after the word onset. The component is observed in the centro-parietal area by EEG (Lau et al., 2008) and the middle temporal gyrus by MEG (Pylkkänen & Marantz, 2003, among others). The N400 is often associated with semantic anomalies. Lau et al. (2008) argue that N400 is considered a reflection of semantic integration. The component has a bigger amplitude when the context is not “semantically-supportive” for both sentential and word targets. For instance, N400 demonstrates a larger amplitude with the presence of an unrelated word before a word target than a related word before a target (e.g., a chair & TEA pair have larger activation than coffee & TEA, the target word is shown in a capital case). In word processing studies, the N400/M350 is a hallmark of lexical processing (Morris et al., 2007; Fruchter & Marantz, 2015).

The N400/M350 consists of several functionally distinct subcomponents (Pylkkänen & Marantz, 2003). In the MEG study, Solomyak and Marantz (2010) demonstrated that lemma (morpheme) frequency and

the TP modulated the M350 in the earlier time window, while the frequency of the whole word did not modulate it. Fruchter and Marantz (2015) also reported that the M350 reflected the semantic well-formedness of complex words in the latter time window. A subsequential study provided a more detailed profile of lexical processing reported from Greek suffixation. Neophytou et al. (2018) reported that after the decomposition into morphemes by the N170, the component has several distinct time stages within the M350: look-up of each morpheme and (re)combination of morphemes. Stockall et al. (2019) replicated this result in English prefixation. They implied that the lexical processing of complex words is a language universal and independent of the orthographical system. These studies demonstrated that the N400/M350 reflect semantic aspects of morphemes within complex words, which are decomposed in the stage of the N170/M170.

3. Experiment 1: Lexical decision in the *hiragana* script

3.1 Background

3.1.1 Introduction

Studies on Japanese de-adjectival nouns offered inconsistent results between two theories: the dual mechanism theory and the single mechanism composition theory. There are two types of nouns derived from the adjective in Japanese (Sugioka, 1986; Kageyama, 1993). The *-sa* noun, “*yowa-sa*,” 弱 - さ , weakness; and the *-mi* noun: “*yowa-mi*,” 弱 - み , weak point (corresponding adjective: “*yowa-i*,” weak). Despite their phonological and morphological transparencies, they have different features in productivity and semantic transparency (Sugioka, 1986). Like English examples, the former *-sa* noun is extremely productive, which can be derived from neologism. In addition, the *-sa* noun is semantically transparent, which is translated by attaching “ness” to the corresponding adjective. (e.g., [X]_{adjective} → [X-ness]_{noun}, being X; [*yowa-i*]_{adjective} → [*yowa-sa*]_{noun}, being weak). The latter *-mi* noun is unproductive in that the nominal suffix *-mi* attaches to only 30 adjectives with monomorphemic *wago*⁷ (Japanese-derived) stems⁸⁹

⁷ Words that are not derived from loan words from Chinese or other languages.

⁸ In this dissertation, we denote derivational root and stem “stem” despite Kageyama (1993:15-19) and other Japanese studies distinguish the morpheme “*yowa*” in “*yowa-mi*” to denote root (*gokon*, 語根), while “*yowa*” in “*yowa-sa*” to denote stem (*gokan*, 語幹) due to their features. However, since this dissertation deals both phonologically and orthographically same forms, such as “*yowa*” shown above, to avoid confusion, we use the term “stem” to denote root and stem.

⁹ Although not mentioned in Kageyama (1993) among others *nama-gusa-mi* (fishy smell) is an exception of *-mi* type for having a bimorphemic stem.

(Sugioka, 1986, pp. 126–146; “*ka-yowa-i*,” frail → “*ka-yowa-sa*,” **“ka-yowa-mi*,” frailness; henceforth referred to as *-mi* type). In addition, due to their semantic opacity (lexicalized meanings), it is difficult to expect their meanings from their corresponding adjectives (e.g., “*yowa-i*,” weak → “*yowa-mi*,” weak point, *weakness). Building on the dual mechanism theory, a neurolinguistic study by Hagiwara et al. (1999) claims that this model can be applied to two types of de-adjectival nouns in Japanese. In their studies, Hagiwara et al. (1999) reported that agrammatic Broca’s aphasic patients showed difficulty in producing *-sa* nouns, with little difficulty in *-mi* nouns, while the other type of aphasic patients, especially the patients of Gogi (word-meaning) aphasia, showed the opposite pattern, i.e., difficulty in producing *-mi* nouns. From these results, they concluded that these differences (or dissimilarities) reflected the two neurologically independent rule-based and memory-based word processing mechanisms, in which the representation in the mental lexicon of *-sa* nouns is a stem and a suffix, while that of *-mi* nouns is a whole word (i.e., not decomposed into morphemes). They claimed that agrammatic aphasics showed impaired processing of *-sa* nouns because they cannot use a grammatical rule to combine a stem and a suffix. Despite their grammatical impairment, the patients could process *-mi* nouns, which are lexicalized and whose meaning can be understood without

requiring the grammatical rule. Subsequential theoretical studies proposed the dual mechanism theory for the two nouns (Ito, 2002; Ito & Sugioka, 2002; Ito & Sugioka, 2016).

However, neurolinguistic studies on healthy people reported evidence for the single mechanism composition theory, in which complex words are processed by a common mechanism regardless of their regularity and productivity. For instance, Clahsen and Ikemoto (2012) reported controversial results against the Dual mechanism theory. They found no significant difference between the two nouns in a lexical decision task, suggesting that both are morphologically decomposed. In their eye-tracking experiment in sentential contexts, on the other hand, they demonstrated that the *-mi* nouns showed longer reading times than the *-sa* nouns. From these results, they concluded that the two nouns had the same decomposable word form representations, and the *-mi* noun required more load for lexical access in sentential context due to their lexicalized meanings. Their result was replicated in another priming study by Fiorentino et al. (2016). However, in contrast to many English studies on word processing, neurophysiological evidence of whether the morphological decomposition occurs has yet to be proposed for the Japanese de-adjectival nouns since the studies are conducted for only behavioral aspects.

Behavioral research on the two nouns implies that two different Japanese writing scripts (i.e., logographic Chinese character, *kanji* & phonetic/moraic character, *hiragana*) assist morphological decomposition (Nakano & Kishimoto, 2019). Therefore, they suggest that the different scripts boundary and the morpheme boundary assist the morphological decomposition (e.g., “*yowa-i*” spelled as “弱_レ” (弱=レ_レ)¹⁰ is matched the boundaries of scripts and morphemes). Thus, the results of previous studies (Clahsen & Ikemoto, 2012; Fiorentino et al., 2016) might be affected by the writing scripts since the boundaries of scripts and morphemes of de-adjectival nouns are matched.

In the current EEG experiment, we investigate the neural mechanism for neurolinguistic processing to solve the inconsistent result in de-adjectival nouns. Specifically, focusing on the N170 (2.2.1), which reflects ortho-morphological decomposition, we examine whether the two types of de-adjectival nouns are decomposed by a lexical decision task. In addition, we use stimuli written in the *hiragana* script¹¹ to exclude unexpected effects in decomposition. The N170 was analyzed as a probe of morphological decomposition for *-sa* and *-mi* nouns. Besides, the N400 was also analyzed

¹⁰ The “=” denotes the boundary of the script between *kanji* and *hiragana*.

¹¹ In Japanese, *kanji* cannot compose derived or inflected words since some morphemes have no orthographic form in *kanji* (e.g., the morphemes “*i*” and “*sa*” have only orthographic form written in *hiragana*)

to investigate the lexicalization of the de-adjectival nouns.

3.1.2 Predictions

Based on recent neurolinguistic and behavioral studies, we test whether the two types of Japanese de-adjectival nouns are processed by common neural mechanisms, especially in an earlier stage, that is, morphological decomposition, which was proposed in previous neuroimaging studies in English past-tense verbs. Under the dual mechanism theory of Hagiwara et al. (1999), the larger amplitude of N170 for *-sa* noun is predicted since the theory assumes no decomposition for the *-mi* noun. On the other hand, under the single mechanism composition theory, the N170 demonstrates no difference between the two de-adjectival nouns since the theory assumes that the two nouns consist of two morphemes. Besides, both theories predict a larger N400 activation and higher processing load (i.e., lower accuracy and longer reaction time) for the lexicalized *-mi* nouns.

3.2 Experiment 1: Lexical decision in *hiragana* script

3.2.1 Participants

Twenty-one native Japanese speakers with normal or corrected-to-normal vision were recruited from Kyushu University in Japan (10 females,

mean age $\pm$ *S.D.*: 22.5 ± 2.0 yrs.). All participants were classified as right-handed based on the Edinburgh handedness inventory (Oldfield, 1971: mean $\pm$ *S.D.*; 89.50 ± 14.59). One participant was excluded from the further analyses due to excessive artifacts in the EEG. All participants provided written informed consent before participating in the study. Approval for the experiments was obtained from the institutional review board of the Faculty of Humanities, Kyushu University.

3.2.2 Stimuli

We selected 25 *-mi* type adjectives¹² from the “Balanced Corpus of Contemporary Written Japanese” (BCCWJ, National Institute for Japanese Language and Linguistics) (Maekawa et al. 2014) based on frequencies of *-mi* noun. As the experimental conditions, the *-sa* noun and *-mi* noun were derived from each adjectival stem (e.g., “*yowa-i*,” weak $\rightarrow$ “*yowasa*,” weakness; “*yowa-mi*,” weak point). As a control condition, we selected 25 non-adjective-derived *wago* nouns that have “*sa*” in the last syllable (ultima) as one control factor (e.g., “*nagisa*,” **“nagi-sa*,” beach). Additional 25 non-adjective-derived *wago* nouns that end with “*mi*” were selected as the other control factor (e.g., “*kagami*,” **“kaga-mi*”: mirror). All control nouns do not

¹² The complete list of adjectives with orthographic representations is contained in the appendix Table S1.

have any pseudo-adjectival form (e.g., **“nagi-i”* and **“kaga-i”*). Due to the small number of monomorphemic words which end with “*sa*” or “*mi*,” sixteen stimuli of the control conditions are bimorphemic compound words (e.g., “*kare-kusa*,” dead grass). Contradictory to the de-adjectival stimuli, all bimorphemic compounds cannot be parsed into ultima and the rest of the word (e.g., **“kareku-sa,”* **“kareku-i”*). Therefore, the four factors are grouped into whether the ultima can be parsed as a morpheme. These stimuli were categorized into two conditions by their derivedness (i.e., Derived or Non-derived nouns) and their ultima (i.e., “*sa*” or “*mi*”). In addition, we created 50 nonce words that have “*sa*” in the ultima and 50 nonce words that have “*mi*” in the ultima (e.g., **“garusa”* and **“garumi”*). None of the nonce words have any pseudo-adjectival form or word form (e.g., **“garu-i”* and **“garui”*). In total, 200 stimuli were used in the experiment, and the ratio between word and nonce word was 50:50. The stimuli were presented in *hiragana* script. In total, 200 stimuli were used in the experiment and the ratio between word and nonce word was 50:50. The stimuli was presented in *hiragana* script.

Following Solomyak and Marantz (2010), we defined the lemma transition probability (TP) between the stem and the suffix for de-adjectival stimuli as follows. Equation (1) is the calculation of the TP, where TP, P,

LFreq, and Freq denote the lemma transition probability, probability, lemma frequency¹³, and surface frequency in the BCCWJ, respectively. For the following analyses, we used log-transformed frequencies and TP to reduce the skewness of the data distribution.

$$TP(stem \rightarrow suffix) = \frac{P(stem - suffix)}{P(stem)} \propto \frac{Freq(stem - suffix)}{LFreq(stem)} \quad (1)$$

For instance, the calculation of the TP (*yowa* → *mi*) is represented in

Equation (2):

$$TP("yowa" \rightarrow "mi") = \frac{P("yowa - mi")}{P("yowa")} \propto \frac{Freq("yowa - mi")}{LFreq("yowa")} \quad (2)$$

¹³ In this dissertation, the lemma frequency was calculated by summing adjective, *-mi* noun de-adjectival verbs (transitive, intransitive, and “-garu” verb), and reduplicated form since the BCCWJ counts these nouns separately (e.g., “*tuyo-i*,” “*tuyo-mi*,” “*tuyo-mar-u*,” “*tuyo-me-ru*,” “*tuyo-gar-u*” (derived from “*tuyo-i*,” strong), and “*taka-daka*” (derived from “*taka-i*,” high); Besides, in BCCWJ, frequency of *-sa* noun is included in the adjectival form, “*tuyo-i*”.

The mean values of these variables across the conditions are shown in Table 1.

Table 1. Summary statistics of experimental stimuli (experiment 1)

Data are shown as mean $\pm$ standard deviation. A summary of two-tailed paired *t*-test between *-sa* noun and *-mi* noun. Note that the paired *t*-tests were not applicable (N. A.) for “Log stem frequency” because the same set of primes was used for both *-sa* and *-mi* nouns, and the length of *-sa* and *-mi* nouns was always equivalent.

Property	Condition		<i>p</i> values
	<i>-sa</i> noun	<i>-mi</i> noun	
Length	3.36 ($\pm$ 0.64)	3.36 ($\pm$ 0.64)	<i>p</i> = .500
Log stem frequency	2.34 ($\pm$ 0.72)		N.A.
Log lemma transition probability (TP)	-1.43 ($\pm$ 0.52)	-1.24 ($\pm$ 0.59)	<i>p</i> = .104

3.2.3 Procedure

The visual lexical decision task was run using PsychoPy 3.0.7, a psychophysics software library written in Python (Peirce, 2007). The displayed fonts were Meiryo and Arial for Japanese and other letters, respectively. Each trial started with a fixation cross (“+”) presented in the center of the screen for 500 ms, followed by a stimulus for 500 ms. The target word remained until the participants’ response and was then replaced by a variable blank serving as an inter-trial interval. The inter-trial interval varied

between 600–1200 ms (Figure 1).

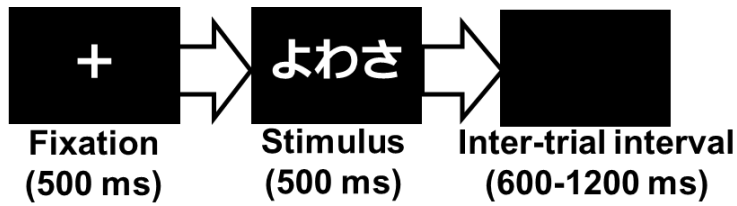


Figure 1. Experimental procedure

The participants performed a visual lexical decision task on the stimulus words by pressing buttons.

During the experiment, the participants were seated in an electromagnetic shield room with a monitor positioned approximately 130 cm in front of them. The participants were instructed to judge whether the displayed letter strings were words by pressing buttons. The lexical decision task was preceded by ten practice trials which included feedback. Each participant took a short break about every five minutes during the experiment.

3.2.4 EEG data acquisition

We acquired the EEG by fitting the participants with an elastic cap with 64 embedded Ag/AgCl electrodes (actiCAP, Brain Products; Neurofax EEG-1200, Nihon Kohden). Two additional Ag/AgCl electrodes were placed below and to the left of the left eye to monitor vertical and horizontal eye movements. The linked earlobes served as a reference. All electrodes' impedances were maintained below 10 k Ω throughout the experiment. The

ERPs were amplified with a bandpass of 1–30 Hz and sampled at 1,000 Hz.

3.2.5 Data analysis

We used analyses of variance (ANOVAs) for behavioral and ERP data analyses. We considered a p -value of less than 0.05 statistically significant in the analyses. We applied the Greenhouse-Geisser correction for all effects involving more than one degree of freedom (Greenhouse & Geisser, 1959, Picton et al., 2000) and corrected the multiple comparisons using Shaffer’s modified sequentially rejective Bonferroni procedure (Shaffer, 1986). For significant main effects with more than three levels, we conducted post-hoc pairwise comparisons using paired t -tests. We also ran simple effects tests for significant interactions. We reported the corrected degrees of freedom, corrected p -values, and effect sizes (generalized eta squared [η^2_G] for ANOVA and Cohen’s d for t -tests; Cohen, 1992).

3.2.6 Data processing

We conducted behavioral data analysis on the trials of the word conditions with corrected answers. In addition, following Baayen (2008), we excluded trials with reaction time exceeding $\pm 2.5 S.D.$

For the processing of the EEG data, we used MNE-Python for the ERP

data analysis (Gramfort et al., 2013). Independent Component Analysis (ICA) was applied to denoise eye blinks and electrocardiograms. ERP averages were computed in a 600 ms time window in all conditions, and the baseline was set to 100 ms before the stimulus onset. Trials with nonce words, incorrect responses, and large artifacts exceeding 80 μ V (peak-to-peak) were excluded from further analyses (approximately 11.3% of the data).

To localize the statistically significant activations of electrodes and time windows, we applied a two-way ANOVA (**Derivedness** [**Derived** and **Non-derived**] $\times$ **Ultima** [**“sa”** and **“mi”**]) and the spatio-temporal permutation cluster analyses for the N170 and the N400 ERP components (Gwilliams et al., 2016). Following Fruchter et al. (2013) and Gwilliams et al. (2016), the time windows of the N170 and the N400 were defined as 100–250 ms and 300–500 ms, respectively. Although the ROIs of the ERP component were not selected a priori, anatomical ROIs (aROIs) were used when the spatial cluster(s) could not be identified. When we found significant main effects or interactions between derivedness and ultima, the spatio-temporal cluster analysis was performed for the amplitude to assess the statistical differences in each condition. In light of the numerous regressions computed in the analysis, we performed a multiple comparisons correction procedure (a permutation test following Maris & Oostenveld, 2007) on the continuous

moving average amplitude of the point-by-point regressions. The estimates (beta values) of fixed effects were computed for 10,000 random permutations to identify temporal clusters in which target variables were significantly correlated with the amplitude.

To examine the influence of individual target variables (i.e., TP, length, and ultima) on the ERP components for the attested cluster(s), we conducted linear mixed-effect (LME) model analyses for the derived condition on the moving average of 5 ms change in millisecond-level amplitude. The lmer function in the lme4 package of R was applied for the analyses (Bates et al., 2009; Bates et al., 2015). To investigate whether the nouns are decomposed into morphemes, we used Equation (3), including the TP between the morphemes for the LME model. As the full model.

$$\mathbf{Amplitude} \sim \mathbf{length} * \mathbf{ultima} * \mathbf{TP} + (1|\mathbf{Stimulus}) \quad (3)$$

The interactions among fixed effects are included in the equation. Following Barr et al. (2013), the maximal random effect was determined as the most complex structure to which the LME models converged at any ROIs. A step function was applied to determine the best model by eliminating fixed effects without significant contribution. We employed Akaike's Information Criterion (AIC) to determine the best model from the full model. In addition,

the simple model without fixed effects was employed as a reference model to compare the goodness of fit among models (Equation (4)).

$$Amplitude \sim (1|Stimulus) \quad (4)$$

Importantly, as the TP is always less than one, log-transformed TP, which was used in linear mixed-effects modeling, is always a negative value (see 3.2.2 for the definition of the TP). In addition, as the N170 and N400 are negative-going potentials, the positive regression coefficient (beta/estimate) in linear mixed-effects models for the TP indicates larger peaks of these components, while negative beta/estimate indicates smaller peaks of these components.

3.3 Results

3.3.1 Behavioral results

Behavioral analyses were conducted for the word conditions. The mean accuracies of the lexical decision task of all conditions were higher than 85 %, which indicated that the participants correctly performed the task.

The two-way ANOVA (**Derivedness** [**Derived** and **Non-derived**] × **Ultima** [**“sa”** and **“mi”**]) for accuracy revealed no significant main effects for derivedness or interaction [**Derivedness**: $F(1, 19) = 0.2568, p = .62, \eta^2_G = .0005$; **interaction**: $F(1, 19) = 1.133, p = .30, \eta^2_G = .0048$]. The main effect

of ultima was marginally significant [$F(1, 19) = 3.182, p = .09, \eta^2_G = .11$] (Table 2). The accuracy for ultima was marginally higher in ultima of “sa”. The two-way ANOVA (**Derivedness** $\times$ **Ultima**) for reaction time revealed no significant main effects [**Derivedness**: $F(1, 19) = 3.298, p = .09, \eta^2_G = .0076$; **Ultima**: $F(1, 19) = 0.6739, p = .42, \eta^2_G = .0003$] (Table 3). However, the interaction was significant [**interaction**: $F(1, 19) = 4.487, p = .048, \eta^2_G = .0041$]. Simple effects for Derivedness $\times$ Ultima demonstrated that derivedness at ultima of “sa” had a significant faster reaction time [$F(1, 19) = 6.113, p = .02, \eta^2_G = .02$]. In addition, Ultima with “sa” at derived condition showed marginally significant faster reaction time [$F(1, 19) = 4.360, p = .05, \eta^2_G = .0069$]. These results suggest that derived condition at the ultima of “sa” requires a significantly lower processing load than the other three factors. The other simple effect was not significant [**Derivedness at Ultima of “mi”**: $F(1, 19) = 0.2105, p = .65, \eta^2_G = .0006$; **Ultima at Non-derived**: $F(1, 19) = 2.068, p = .17, \eta^2_G = .0021$].

Table 2. ANOVA results for accuracy (experiment 1)

df, degrees of freedom; η^2_G , generalized eta squared.

Factors	<i>df</i>	<i>F</i>	<i>p</i>	η^2_G
Derivedness $\times$ Ultima				
Derivedness	1, 19	0.2568	.62	.0005
Ultima	1, 19	3.182	.09	.11
Interaction	1, 19	1.133	.11	.0048

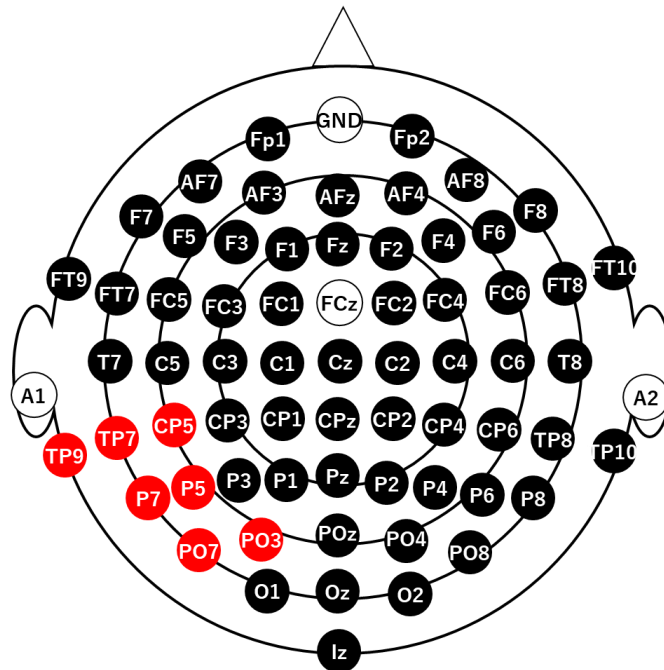
Table 3. ANOVA results for reaction times (experiment 1)*df*, degrees of freedom; η^2_G , generalized eta squared.

Factors	<i>df</i>	<i>F</i>	<i>p</i>	η^2_G
Derivedness $\times$ Ultima				
Derivedness	1, 19	3.298	.085	.0076
Ultima	1, 19	0.6739	.42	.0003
Interaction	1, 19	4.487	.048	.0041

3.3.2 ERP results

3.3.2.1 N170: Morphological decomposition

For the analysis of the N170, the spatio-temporal analysis attested spatial cluster, which covers almost all electrodes. Therefore, aROI was used for spatio-temporal analysis of the N170 (Figure 2). The aROI was determined by the previous MEG studies on the corresponding MEG

**Figure 2.** aROIs of the N170

Red electrodes are dealt with the aROI of the N170.

component of the N170 (Gwilliams et al., 2016).

The two-way ANOVA (**Derivedness** $\times$ **Ultima**) revealed the significant main effect for derivedness in one cluster, which covers all electrodes of the aROI ($p = .0175$). Moreover, the derived words have a larger N170 (Figure 3). The significant temporal cluster was 100–207 ms. On the other hand, the main effect of ultima and interaction did not have any significant temporal clusters ($ps > .2$).

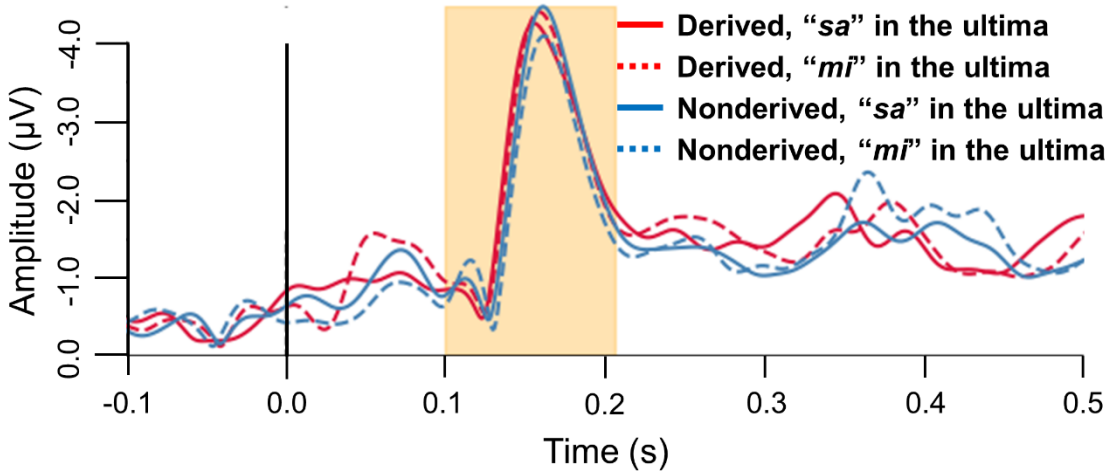


Figure 3. Grand averages of the N170 among the four factors. The Derivedness factors showed significant larger amplitude ($p = .0175$). The highlighted area is significant temporal cluster for main effect of Derivedness. Red and blue lines denote Derived and Underived conditions, respectively. Solid and broken lines denote “sa” and “mi” Ultima, respectively.

The following LME analysis demonstrated that the best LME model was the one including the TP as the fixed effect and stimulus, as illustrated in Equation (5):

$$Amplitude \sim TP + (1|Stimulus) \quad (5)$$

Compared to the simple and full models, the best model had a lower AIC than the other models (Table 4). This result indicated that the TP modulated the N170, not the ultima of each stimulus. The derived condition is decomposed, whereas the non-derived condition is not. In addition, the TP had significant positive correlations with the N170 in one temporal cluster (Figure 4: 143–158 ms).

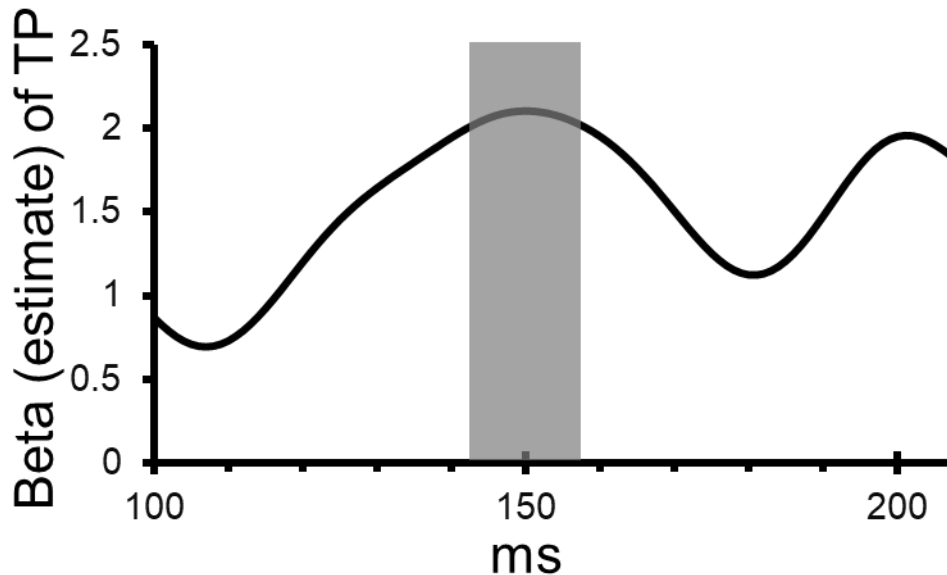


Figure 4. Beta (estimate) of TP for the Best LME model
The highlighted area (143–158 ms) represents the corrected $p < .05$

Table 4. The comparison among LME models

AIC, Akaike's information criterion; df , degrees of freedom

Models	No. params	AIC	Log-Likelihood	χ^2	df	p value
Simple	3	10863	-5428.7			
Full	10	10868	-5424.0	5.3	6	0.51
Best	4	10861	-5426.6	4.2	1	0.041

3.3.2.2 N400: Lexical processing

For the analysis of the N400, the spatio-temporal analysis also attested spatial cluster, which covers almost all electrodes. Therefore, aROI was used for spatio-temporal analysis of the N400 (Figure 5). The aROI was determined by the previous EEG studies of the N400 (e.g., Lavric et al., 2007).

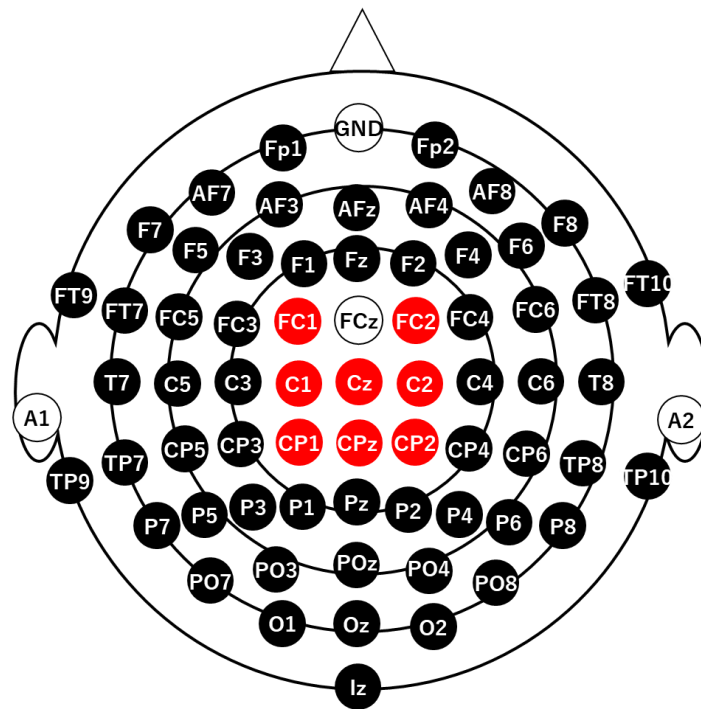


Figure 5. aROIs of the N400
Red electrodes are dealt with the aROI of the N400.

However, the two-way ANOVA (**Derivedness** $\times$ **Ultima**) revealed any significant main effects or interaction ($ps > .1$). Furthermore, these analyses for the N400 had large negative-going components, which differ from the typical N400 components, peaks around 550 ms from the onset of the stimuli

(Figure 6). Further analyses for the N400 were not conducted due to the lack of significant clusters.

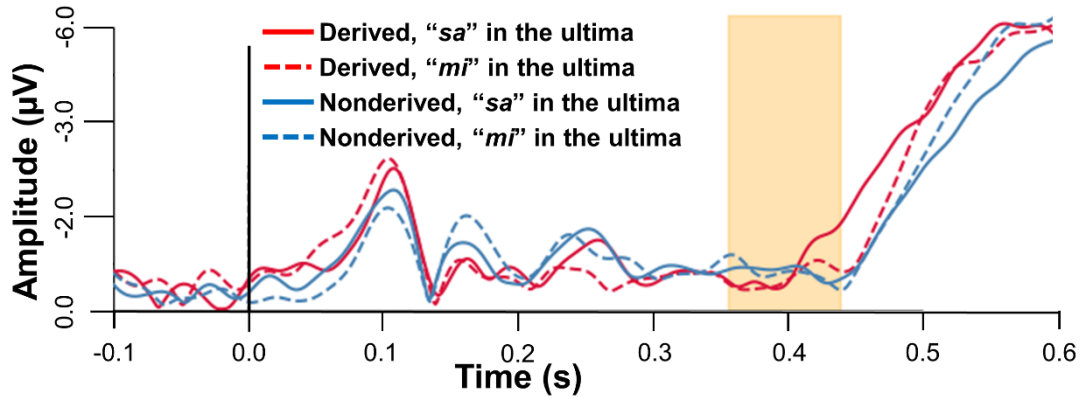


Figure 6. Grand averages of the N400 among the four factors. The four factors did not show significant differences ($ps > .1$). The highlighted area is temporal cluster for main effect of Derivedness found by the permutation ($p = .1470$). Red and blue lines denote Derived and Underived conditions, respectively. Solid and broken lines denote “*sa*” and “*mi*” Ultima, respectively.

3.4 Discussion

In the current experiment using phonetic/moraic *hiragana* as a stimulus, we obtained four significant remarkable results by examining the behavioral and neurophysiological results. First, in the analysis of the N170, we demonstrated a significant main effect for the derived condition (Figure 3). In addition, the amplitudes of the N170 were significantly larger in the derived condition. These results suggest that the two types of de-adjectival derived nouns are decomposed similarly, whereas the other non-derived nouns are not. This supports the single mechanism composition theory.

Second, in comparison among the linear mixed-effect models, we obtained the positive correlation between the fixed effect TP and the amplitude of N170, while the other fixed effects and interactions do not (Figure 4 & Table 4). This further proves that the derived condition is decomposed into morphemes (i.e., stem and suffix). In addition, the positive correlation indicates that the larger TP enhances the amplitude of N170. Third, we did not obtain any clusters for the N400, which reflect lexical processing (Figure 6). The ERPs for all factors showed atypical components for the N400. Fourth, we found a significant difference in the interaction between derivedness and ultima in the reaction time analysis. This implies that the *-mi* noun requires a larger processing load than the *-sa* noun.

In the two-way ANOVA analysis for the N170, we found a significantly lower amplitude for the derived condition than the non-derived condition (Figure 3; $p = .0175$). As described in 2.2.1., the negative-going component N170 (and M170) evokes larger activation for complex words than simple ones. This result demonstrated that the two types of de-adjectival nouns are decomposed (i.e., recognized as complex words), whereas the non-derived condition is not. The result replicates the previous behavioral study for healthy people and is consistent with the single mechanism composition theory (Clahsen & Ikemoto, 2012; Fiorentino et al., 2016). On the other hand,

the dual mechanism theory, which assumes *-mi* noun is memorized as a whole word, cannot explain this result.

In comparing the LME models for the N170, we found that the model that includes only the TP as a fixed effect is the best model (Table 4). This result further supports the single mechanism composition theory since this theory predicts the TP as a fixed effect. On the contrary, the result also provides negative evidence for the dual mechanism theory, as the theory predicts an interaction between the TP and *ultima* as a fixed effect. Moreover, the time window of significant correlation between the TP and the N170 amplitude (Figure 4: 143–158 ms) replicates the other studies of the single mechanism composition theory for morphologically complex words using N/M170 (Table 5). It should also be noted that the result provides further evidence for morphological typology in Japanese and other typologically different languages.

Moreover, this result demonstrated that matching the boundaries of scripts and morphemes does not affect the morphological decomposition since the current experiment using only the *hiragana* script replicates the result of the other studies on the N/M170. Therefore, the derived nouns are decomposed regardless of their script. Meanwhile, the result of the LME model had a positive regression coefficient (beta/estimate) for the TP. In

other words, as a larger N170 peak reflects a higher processing load for decomposition, this indicates that when the TP of stimulus is smaller, the following suffix becomes more difficult to predict. However, this result is difficult to interpret since the more difficult-to-predict succeeding suffix, the N170, which reflects the processing load, decreases.

Table 5. A summary of neurolinguistic studies demonstrating the correlation between N/M170 and TP of affixes.

Study	Language	Type of word formation (lexical category)	Morphological type	Latency (ms)	Morphological variable
Current experiment	Japanese	derivation (noun)	suffixation	143-158	stem:suffix TP
Ohta, Oseki & Marantz (2019)	Japanese	Inflection (verb)	suffixation	150-200	root:suffix TP
Zweig & Pykkänen (2009)	English	derivation (noun and verb)	prefixation suffixation	prefix: 174-182 suffix: 170-186	Complexity (simple or bimorphemic)
Solomyak & Marantz (2010)	English	derivation (noun)	suffixation	178-214	stem: whole word TP
Lewis, Solomyak & Marantz (2011)	English	derivation (noun)	suffixation	164-208	stem: whole word TP
Fruchter, Stockall & Marantz (2013)	English	inflection (irregular verb)	irregular	146-196	morphophonological congruency
Gwilliams & Marantz (2018)	English	derivation (noun)	suffixation	150-180	stem: whole word TP
Stockall et al., (2019)	English	derivation (verb and noun)	prefixation	100-200	stem: whole word TP
Neophytou et al., (2018)	Greek	derivation (noun)	suffixation	100-200	stem: whole word TP
Kastner, Pykkänen & Marantz (2018)	Hebrew	inflection (root and template: verb, noun, and adverb)	transfixation	177-219	Prosodic match between root and template
Wray, Stockall & Marantz (2022)	Tagalog	derivation (noun and verb, not differentiated)	infixation, circumfixation, and reduplication	150-200	stem: whole word TP
Vartiainen et al., (2009)	Finnish	inflection (noun)	suffixation	200-800	complexity
Hakala et al., (2018)	Finnish	derivation and inflection (noun)	suffixation	140-200	Morfessor minimum description length

In the two-way ANOVA analysis for the N400 time window, we did not find any cluster among the four factors (Figure 6). In addition, the ERP components peaks around 550 ms after the stimulus onset; this is atypical for the N400, which peaks around 350 ms (Lau et al., 2008; Lavric et al., 2007). Although it is difficult to interpret, this result implies that the task of the current experiment required a larger and unexpected load for lexical processing (i.e., lexical access and composition). Additionally, some participants reported difficulties in the lexical decision that the stimuli were presented in the less common *hiragana* script. Intriguingly, Nakano et al. (2016) reported that the presentation of a single *kanji* activates the meaning of words and orthography in the mental lexicon. Therefore, presenting Japanese stimuli only in *hiragana* may cause an unexpected effect on the time windows for the lexical processing. Thus, whether the result of the N400 supports which theory is unclear.

In the reaction time analysis, the interaction between derivedness and ultima was significant. Moreover, simple effects for derivedness at the ultima of “*sa*” demonstrated a significantly shorter reaction time than the other conditions. These results suggest that a difference in semantic transparency of de-adjectival nouns can be observed at the behavioral level. This interpretation can explain the behavioral dissociation of *-sa* and *-mi* nouns

(Hagiwara et al., 1999) and longer reading times for *-mi* nouns in a sentential context (Clahsen & Ikemoto, 2012). These results demonstrate that a higher processing load for *-mi* noun is predicted by the two theories of word processing.

The current experiment demonstrated the inconsistent result to the previous study, which support the dual mechanism theory in the decomposition process (Hagiwara et al., 1999). However, the current experiment supported the studies of the single mechanism composition theory (Clahsen & Ikemoto, 2012; Fiorentino et al., 2016). Meanwhile, the behavioral result replicated the difficulty in processing the lexicalized *-mi* noun, which is predicted similarly by both theories. Differences in the experimental paradigms may cause these inconsistent results. Namely, Hagiwara et al. (1999) required production for aphasic patients, while the other studies required comprehension for healthy adults.

In this experiment, the decomposition was observed by the stimuli written in the *hiragana* script, which indicated that the matching between the morpheme and script boundaries does not affect the automatic morphological decomposition. This result is consistent with Ohta et al. (2019), which demonstrated that the decomposition in Japanese is dependent on the morpheme boundary but independent of the letter boundary. These

results imply that the morphological decomposition in visual modality is driven by the abstract morpheme but not by the letters in the language with mixed scripts, such as Japanese. Therefore, we can conclude that the coincidence or mismatch between the scripts/letters and morpheme boundary does not affect the N170. This supports the prediction of the single mechanism composition theory. In the analysis of the N400, however, there were no significant differences among the four conditions. In addition, the ERP component in the time window of N400 did not have the typical N400, which peaks around 350–400 ms after the onset but larger peaks around 550 ms after the onset. This result implied that the N400, the lexical processing, needed more load when written in the less common script (i.e., *hiragana* only).

To summarize the results of the EEG (N170 and N400), the presentation only in the *hiragana* script reflects brain activation's automatic decomposition (N170), which correlates with TP. In contrast, the lexical processing (N400) does not reflect brain activation because of the less common orthographical presentation (i.e., *hiragana*). This difference between the N170 and N400 in the current experiment suggests that the writing system with mixed scripts, such as logographic kanji and phonetic

(moraic) *hiragana*, affect the two stages of processing complex words differently: the automatic decomposition, N170, are not affected by the script, while the lexical processing, N400, are affected by the script. Although similar writing systems with mixed scripts like Japanese can be seen in other languages, all of these had extinct. (Egyptian, Akkadian, Sumerian: Coulmas, 2003; Hieroglyphic Hittite: Melchert, 2003; Hieroglyphic Luwian: Payne, 2014). For this reason, the Japanese writing system with a mixed script is unique among the existing languages. Therefore, more detailed studies are needed to interpret the difference in word processing between the scripts.

The behavioral results on the reaction time demonstrated and replicated the higher processing load for the *-mi* noun than the *-sa* noun. Although the uncommon *hiragana* script might affect the reaction time, the accuracy result suggests that the participants performed the lexical decision task collectedly.

The current neurophysiological experiment supported the single mechanism composition theory for morphological decomposition (i.e., N170) but not the dual mechanism theory. However, the result of the latter stage of processing (i.e., N400: lexical processing) was difficult to interpret due to its atypical component. The current result demonstrates the single

mechanism composition theory for the two de-adjectival nouns in morphological decomposition. However, as reviewed in 2.2.1., the correlation between N/M170 and the TP reflects garden path-like morpho-orthographical decomposition for genuine and pseudo-complex words (e.g., Rastle et al., 2004; Lewis et al., 2011; e.g., “farm-er” and “broth-er”). Therefore, there is a possibility that the result of N170 reflects the decomposition of pseudo-complex words for the *-mi* noun. To solve the inconsistent results of word processing in the de-adjectival noun, further analysis for the lexical processing, N400, is needed.

In Chapter 4, we investigate the neural profile of the N400. Specifically, we investigate whether the amplitude of N400 correlates with the TP. In addition, to investigate the N400, another experimental paradigm, masked priming is introduced in the next chapter.

3.5 Conclusion

In the current experiment, we aimed to fill the gap of the contradiction of studies in Japanese de-adjectival nouns. The two contradictory theories have been proposed for the *-sa* and *-mi* nouns; the dual mechanism theory, which assumes no decomposition for the *-mi* noun, and the single mechanism composition theory, which assumes decomposition for the two de-adjectival nouns. Using the ERP N170 as a hallmark of decomposition, we examined which model is plausible to solve the contradiction. In the current ERP analysis, the N170 components demonstrated a common decompositional mechanism, and the activation is correlated with the TP. Therefore, the N170 result was consistent with the single mechanism composition theory, such as English. In particular, by applying LME models, we elucidated that the TP of nouns modulates the N170 component, while the whole word frequency does not.

Furthermore, the modeling demonstrated that the N170 was positively correlated with the transition probability between morphemes (Figure 4) and their correlated time window (Table 5). Therefore, the results of the N170 supported the single mechanism composition theory. However, there are two problems with the current experiment. First, as reviewed in 2.2.1, whether the two de-adjectival nouns genuinely dealt with their morphemic units in

word processing is debatable since the N170 also reflects orthographical decomposition. In contrast to the N170, the result of N400, the lexical processing was uninterpretable. The results of the current experiment suggested that the presentation of stimuli in the *hiragana* script affected differently between the N170 and N400 and caused unexpected effects for the N400 component. For this reason, further evidence is needed to examine a more detailed mechanism for word processing, specifically for the component of lexical processing, N400.

To solve these problems, in the next chapter, we examine the component N400, which reflects lexical processing. Using the masked priming paradigm, we elucidate whether the decomposed morphemes of the de-adjectival nouns are genuinely dealt with morphemes in the latter process of complex words. In addition, we use mixed scripts (*kanji* and *hiragana*) to avoid unexpected effects on the N400.

4. Experiment 2: Masked priming

4.1 Background

4.1.1 Introduction

Experiment 1 demonstrated that the single mechanism composition theory outperformed the other dual mechanism model, strongly suggesting that morphological decomposition occurs in the *-mi* noun replicating Clahsen and Ikemoto (2012) and Fiorentino et al. (2016). However, there were two problems with experiment 1. First, whether the *-mi* nouns genuinely dealt with their morphemic units is unclear as the decomposition occurs not only morphologically but orthographically (2.2.2.: “farm-er” and “broth-er”; Rastle et al., 2004; Lewis et al., 2011; Lavric et al., 2007). Second, the result of the N400 is uninterpretable, probably due to the less common script presented only in the *hiragana*. Therefore, to solve the contradiction of Japanese de-adjectival word processing (Introduction aim (1)), we need to further investigate the abovementioned problems, whether the decomposition is morphological, not orthographical.

In this chapter, focusing on another neural activation, N400, which reflects lexical processing, we further examine the contradiction of Japanese de-adjectival nouns. Specifically, adapting the masked priming paradigm, which affects reaction time and the N400, we investigate the remained problems of the Japanese de-adjectival nouns. The detail of the masked

priming paradigm is explained in the next section. For further analysis, applying the LME models to estimate the amplitude of the N400 time window, we elucidate whether or how the target stimulus's statistical variables modulate the amplitude of the N400. In addition, the N170 component was also analyzed to examine its Laterality. Besides, to avoid unexpected effects on the N400, most primes and targets are shown in the mixed script (i.e., *hiragana* and *kanji*).

4.1.2 Experimental paradigm: Masked priming

The masked priming paradigm is widely applied to investigate the visual processing of morphologically complex words. In this paradigm, a word stimulus called prime is presented in a short time (< 50 ms) after preceding symbols (e.g., “#####”), which acts as a forward mask, followed by a word called target, which acts as a backward mask (Forster & Davis, 1984). Due to the forward mask, the target mask, and the short presentation time of the prime, participants do not consciously recognize what is presented as the prime. This paradigm has been used to elucidate how language processing is reflected in behavior and temporal/spatial neural activities by applying neurolinguistic techniques such as EEG, MEG, and fMRI.

Using this paradigm, Rastle et al. (2000) reported that the facilitation of reaction time called the “priming effect” for morphologically related¹⁴ prime and target (e.g., teacher & TEACH)¹⁵, while semantically related pairs did not (e.g., cello & VIOLIN). This result indicates that the prime of morphologically related pair lightens the processing load of the target due to the shared morpheme of these pairs being activated through priming. Intriguingly, similar to the N170/M170 amplitude mentioned in 2.2.1., this facilitation is also observed for pseudo-complex (exhaustively parsable) pair (e.g., brother & BROTH), but not for orthographically genuine controls (e.g., brothel & BROTH) nor orthographically similar pairs (e.g., boil & BROIL) (Rastle et al., 2004), suggesting that the masked priming paradigm can provide a situation where morphological effects are observed “in the absence of semantic effects” (Rastle et al. 2000: 517). Furthermore, Crepaldi et al. (2010) reported that irregular verb pairs (e.g., fell & FALL, mice & MOUCE) had a similar facilitating priming effect for regular verb pairs, while pseudo-irregular pairs (e.g., bell & BALL, dice & DOUSE) did not. These behavioral studies suggest that morphologically complex words are processed by primarily using morphological cues (i.e., identical morphemes

¹⁴ The term “(morphologically) related” indicates that a word in the prime target pairs is able to be parsed into existing morphemes regardless of their meanings.

¹⁵ In the masked priming study, the prime and target are usually presented in lower and capital case, respectively.

between the prime and target). Orthographic cues (e.g., similarity of the letters) also play a role in the masked priming paradigm, but it should be noted that orthographic cues alone cannot elicit the robust priming effect (e.g., *brothel* & *BROTH*). In summary, these behavioral studies suggest that morphologically complex words are processed by morpho-orthographical rules, which also include abstract morphemes of a word.

In addition, the masked priming paradigm also affects the neural activation, N400/M350 (lexical processing) and N170/M170 (morphological decomposition). In an ERP study employing masked priming, Lavric et al. (2007) reported that morphologically related prime-target pairs attenuated the peak of N400 in the parieto-occipital area, while orthographically related prime & target pairs did not. In an MEG study employing the masked priming paradigm, Fruchter et al. (2013) also reported a difference in the M350 neural activation between morphologically related and unrelated pairs. The attenuation is also called the “priming effect.” These studies demonstrate that morpho-orthographically related prime-target pairs lighten the processing load of the target due to the shared morpheme of these pairs being activated through priming. Hence, employing the masked priming paradigm in ERP studies is an effective way to investigate the processing of morphologically complex words.

In an MEG study, Lewis et al. (2011) also demonstrated that only morpho-orthographically complex words enhance the amplitude of the M170, while the orthographically related control does not (brother vs. brothel). In addition, using the masked priming paradigm, Fruchter et al. (2013) reported that the masked priming effects on M170 activity were modulated by morpho-orthographically complex pairs and irregular verb pairs but not by pseudo-irregular pairs (jumped & JUMP, fell & FALL vs. bell & BALL). Taken together, these results demonstrate that the masked priming paradigm and M170 allow us to reveal the decomposition of morpho-orthographic properties for complex words (see also Gwilliams 2020).

Therefore, applying the masked priming paradigm, which modulates the reaction time and neural activations, allows us to estimate the modulation of the lexical processing, which remains unclear in experiment 1.

4.1.3 Predictions

In this experiment, we employ the masked priming paradigm to test the neural mechanism of morpho-orthographic and semantic aspects by behavioral and electrophysiological evidence. Based on recent neurolinguistic and behavioral studies, we test the single mechanism

composition theory that the two types of Japanese de-adjectival nouns are processed by common neural mechanisms, especially in a later stage (i.e., N400: lexical processing). Under this hypothesis, the N400 will be attenuated when the prime-target pairs have common stems (i.e., priming effects on the N400), and behavioral data for these de-adjectival nouns will show significant priming effects. Moreover, we conduct the linear mixed effect model to test whether the transition probability between morphemes or whole word frequency can estimate the amplitude of N400. Under this hypothesis, the transition probability estimates the amplitude of N400 while the whole word frequency does not. In addition, due to their lexicalized meanings, we further hypothesize that *-mi* nouns will show a larger N400 peak than *-sa* nouns in the later stage, which reflects lexical access and recombination of morphemes. Besides, the left-lateralized N170 will be observed in both de-adjectival nouns. Moreover, behavioral data for these de-adjectival nouns will show significant priming effects, and the N400 will be attenuated when the prime-target pairs have common stems (i.e., priming effects on the N400).

In the current experiment, we focused on four prime and target conditions of the *-mi* type (**Prime type** × **Suffix**) to elucidate the processing of the de-adjectival nouns: **Related condition of *-sa* nouns** (e.g., “*yowa-i*”

→ “*yowa-sa*,” weak → weakness), **Unrelated condition of -mi nouns** (e.g., “*tura-i*” → “*yowa-sa*,” painful → weakness), **Related condition of -mi nouns** (e.g., “*yowa-i*” → “*yowa-mi*,” weak → weak point), and **Unrelated condition of -mi nouns** (e.g., “*tura-i*” → “*yowa-mi*,” painful → weakness). Generally, the more complex or derived word is used as a prime for the prime-target pair in the masked prime paradigm (e.g., teacher & TEACH). In this study, however, we used derived word as prime because we aim to investigate the behavior and EEG brain activity of two types of the de-adjectival *-sa* and *-mi* nouns, not the adjective. In addition, additional de-adjectival nouns that cannot derive *-mi* nouns (non-*mi* types: e.g., “*tura-sa*” and “*hoso-naga-sa*,” painful and slender) are analyzed to examine whether or how the complexity of words affects behavioral and ERP data.

4.2 Experiment 2: Masked priming

4.2.1 Participants

Nineteen native Japanese speakers with normal or corrected-to-normal vision were recruited from Kyushu University in Japan (12 females, mean age $\pm$ *S.D.*: 23.4 ± 2.3 yrs.). All participants were classified as right-handed based on the Edinburgh handedness inventory (Oldfield, 1971: mean $\pm$ *S.D.*; 88.89 ± 13.56). One participant was excluded from further analysis because

of excessive artifacts in the EEG data. All participants provided written informed consent before participating in the study. Approval for the experiments was obtained from the institutional review board of the Faculty of Humanities, Kyushu University.

4.2.2 Stimuli

Based on adjectival stem frequencies, we selected 78 prime words¹⁶ from BCCWJ. We created 156 target words by replacing the last syllable with *-sa* or *-mi*. The prime-target pairs were categorized into three groups characterized by their stems (26 pairs each): *-mi* type (with a simple stem), non-*mi* type with a simple stem, and non-*mi* type with a complex stem. The *-mi* type condition allows two types of derivation (e.g., “*yowa-i*” → “*yowa-sa*,” “*yowa-mi*,” weak → weakness, weak point), while the non-*mi* types do not allow *-mi* suffixation (e.g., “*tura-i*” → “*tura-sa*”/*“*tura-mi*,” painful → painfulness; “*hoso-naga-i*” → “*hoso-naga-sa*”/*“*hoso-naga-mi*,” long and thin → slenderness). We created 156 nonce words by changing the last syllable(s) of the adjectival stem “*-i*” to “*-sii*,” or “*si-i*” to “*-i*” (e.g., “*yowa-i*,” weak → **“yowasi-i*,” “*kurusi-i*,” distressed → **“kuru-i*”). We presented these stimuli (primes and targets) written in a combination of *kanji*

¹⁶ The complete list of adjectives with orthographic representations is contained in appendix Table S2.

(logographic Chinese characters) and *hiragana* (Japanese syllabic/moraic characters), except for the following cases, which may elicit unexpected orthographic effects (see Appendix for the list of the adjectives used). First, to avoid the ambiguity of reading, the word “辛さ,” which could be read as “*kara-sa*” (spiciness) or “*tura-sa*” (painfulness) was represented as “からさ,” “つらさ.” In addition, the word, “鈍さ,” is represented as “とろさ” since it could be read as “*toro-sa*” (stupidity/slowness) or “*nibu-sa*” (dullness/slowness).

Second, phonetic equivalents, kanji used as phonetic symbols rather than for their meanings, were also presented in *hiragana*. For example, “面白さ”¹⁷ was shown as “おもしろさ” (“*omosiro-sa*,” interest). The same things applied to the following words: “おかしさ” (“可笑しさ,” “*okasi-sa*,” fun), “かわいさ” (“可愛さ,” “*kawai-sa*,” cuteness), and “気まずさ” (“気不味さ,” “*kimazu-sa*,” unpleasantness). In addition, the words whose corresponding nonce words can be read as existing Japanese words were displayed in *hiragana*. For example, the corresponding nonce words of “苦しさ” was represented as “苦さ,” which can be read as “*niga-sa*” (bitterness), so this word was represented in *hiragana* (“くるしさ,” “*kurusi-sa*,” distress).

¹⁷ Two kanji, “面” and “白,” represent face/mask and white, respectively.

The same things applied to the following words: “にがさ” (苦さ, “*niga-sa*,” bitterness), “難しさ” (難しさ, “*muzukasi-sa*,” difficulty), and “わるさ” (悪さ, “*waru-sa*,” badness). In addition, the unrelated condition of prime & target pairs which have morphological relationships were excluded from the stimuli (e.g., 早い and 素早さ, “*haya-i*” and “*su-baya-sa*,” early-quickness).

To examine the effect of word form and lexical properties on brain activity, the current expected compared the effects of length, surface frequency, and TP on ERP components as variables. The length of a word is defined as the orthographic presentation of the length of letters. The mean values of these variables across the three conditions, and analysis of variance (ANOVA) statistical test data, are illustrated in Table 6. The TP (log lemma transition probability) was defined in the same manner of experiment 1 (3.2.2.).

Table 6. Summary statistics of experimental stimuli (experiment 2)

The data are illustrated as mean $\pm$ standard deviation. A summary of two-tailed paired *t*-test between *-sa* nouns and *-mi* nouns. Note that the paired *t*-tests were not applicable (N. A.) for “log frequency (prime)” and “length (target)” because the same set of primes was used for both *-sa* and *-mi* nouns and the length of *-sa* and *-mi* nouns was always equivalent.

Property	Condition		<i>t</i> values	<i>p</i> values
	<i>-sa</i> noun (<i>-mi</i> type)	<i>-mi</i> noun (<i>-mi</i> type)		
Log frequency (Prime)	3.57 ($\pm$ 0.64)		N. A.	N.A.
Length (Target)	2.54 ($\pm$ 0.80)	2.54 ($\pm$ 0.80)	N. A.	<i>p</i> = .500
Log stem frequency (Target)	2.26 ($\pm$ 0.08)		<i>t</i> (25) = 2.72	N.A.
Log transition probability TP: (Target)	-1.44 ($\pm$ 0.26)	-1.75 ($\pm$ 0.22)	<i>t</i> (25) = 2.72	<i>p</i> = .104

4.2.3 Procedure

We conducted a visual lexical decision task with masked priming, using PsychoPy 3.0.7, a psychophysics software program in Python (Peirce, 2007). The displayed fonts were Meiryo and Arial for the Japanese and other letters, respectively. Each trial started with a fixation cross (“+”) presented in the center of the screen for 500 ms, followed by a blank for 500 ms. Following the blank, a string of hash marks for 500 ms (“#####”) was presented, which served as a forward mask for the incoming prime word (Figure 7). The adjective prime word (e.g., 弱い, “*yowa-i*,” weak) was presented after the string of hash marks and remained for 48 ms; it was followed by the target word (e.g., 弱さ, “*yowa-sa*,” weakness: 弱み, “*yowa-mi*,” weak point) about which the participant was required to make a lexical decision. The target word was retained until the participant responded and replaced by a variable blank serving as an inter-trial interval. The inter-trial interval varied between 600–1200 ms. The primes and targets were presented randomly for each participant (Figure 7).

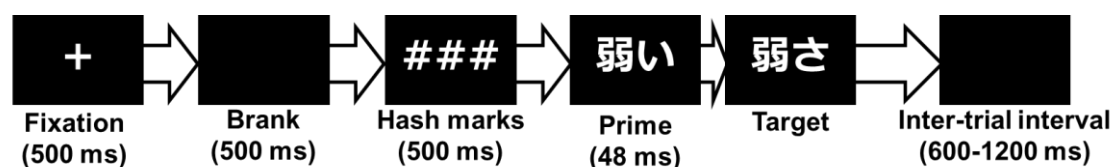


Figure 7. Experimental procedure using the masked priming paradigm
The participants performed a visual lexical decision task on the stimulus words by pressing buttons.

The figure illustrated a related condition of the *-sa* noun.

During the experiment, the participants were seated in an electromagnetic shield room with a monitor positioned approximately 130 cm in front of them. Participants were instructed to judge whether the displayed letter strings were words by pressing buttons. The lexical decision task was preceded by eighteen practice trials which included feedback. Each participant took a short break about every five minutes during the experiment.

4.2.4 EEG data acquisition

We acquired the EEG by fitting the participants with an elastic cap with 64 embedded Ag/AgCl electrodes (actiCAP, Brain Products; Neurofax EEG-1200, Nihon Kohden). Two additional Ag/AgCl electrodes were placed below and to the left of the left eye to monitor vertical and horizontal eye movements. The linked earlobes served as a reference. All electrodes' impedances were maintained below 45 k Ω throughout the experiment. The ERPs were amplified with a bandpass of 1–30 Hz and sampled at 1,000 Hz.

4.2.5 Data analysis

We used repeated-measures analyses of variance (rANOVAs) for behavioral and ERP data analyses. We considered a *p*-value of less than 0.05 statistically significant in the analyses. We applied the Greenhouse-Geisser

correction for all effects involving more than one degree of freedom (Greenhouse & Geisser, 1959, Picton et al., 2000) and corrected the multiple comparisons using Shaffer's modified sequentially rejective Bonferroni procedure (Shaffer, 1986). For significant main effects with more than three levels, we conducted post-hoc pairwise comparisons using paired *t*-tests. We also ran simple effects tests for significant interactions. We reported the corrected degrees of freedom, corrected *p* values, and effect sizes (generalized eta squared [η^2_G] for rANOVA and Cohen's *d* for *t*-tests, Cohen, 1992).

4.2.6 Data processing

We excluded trials with incorrect responses and results exceeding ± 2.5 *S.D.* from the average (approximately 8.4% of the data) (Baayen, 2008). We applied this exclusion procedure to the behavioral data. We conducted two independent analyses for the accuracy and reaction times: We ran a two-way ANOVA ([**Related** and **Unrelated**] $\times$ **Suffix** [**-sa** and **-mi**] for **-mi** type (e.g., “yowa-sa,” “yowa-mi”) to examine the priming effect of different de-adjectival nouns, and an additional two-way rANOVA (**Prime type** [**Related** and **Unrelated**] $\times$ **Stem type** [**non-mi** type with a simple stem, and **non-mi** type with a complex stem]) for **-sa** nouns (e.g., “yowa-sa,” “tura-sa,” and

“*hoso-naga-sa*,” weakness, stressfulness, and, long and thin, respectively) to compare the priming effect of different stem types.

For the ERP data analysis, we used MNE-Python (Gramfort et al., 2013) and applied an independent component analysis (ICA) to de-noise the eyeblinks and ECGs. We computed ERP averages for a 600 ms time window under all conditions, and the baseline was set to 100 ms before the stimulus onset. We excluded trials with incorrect responses and large artifacts exceeding 80 μ V from further analysis (approximately 11.8% of the data).

For the analyses of the N400, epoched ERPs were averaged for the following seven regions of interest (**ROIs**): front central (**FC**: Fp1, Fp2, AF3, AFz, AF4, F1, Fz, F2), left frontal (**LF**: AF7, F7, F5, F3, FT9, FT7, FC5, FC3), central (**C**: FC1, FC2, C1, Cz, C2, CP1, CPz, CP2), right frontal (**RF**: AF8, F8, F6, F4, FT10, FT8, FC6, FC4), left temporal (**LT**: T7, C5, C3, TP9, TP7, CP5, CP3, P7, P5, P3), right temporal (**RT**: T8, C6, C4, TP10, TP8, CP6, CP4, P8, P6, P4), and parieto-occipital (**PO**: P1, Pz, P2, PO7, PO3, Pz, PO4, PO8, O1, Oz, O2, Iz) (Figure 8A). Following other masked priming studies of the N400/M350, the time window for the N400 was restricted to 300–500 ms (Morris & Stockall, 2012; Fruchter, et al., 2013). For the *-mi* type (*-sa* and *-mi* nouns), to identify the ROIs in which the N400 differed between related and unrelated conditions, we ran a three-way rANOVA

(**Prime Type** [**Related** and **Unrelated**] $\times$ **Suffix** [*-sa* and *-mi*] $\times$ **ROIs** [FC, LF, C, RF, LT, RT, and PO]) on the mean amplitude of the time window of the N400. When significant main effects or interactions were found, we performed follow-up two-way rANOVA to assess the presence of priming effects in each condition.

For the non-*mi* types (*-sa* nouns), we ran a two-way rANOVA (**Prime Type** [**Related** and **Unrelated**] $\times$ **ROIs**) on the mean amplitude of the N400. Based on the three-way rANOVA on the *-mi* type, the identified ROIs in which the N400 differed between related and unrelated conditions are not analyzed.

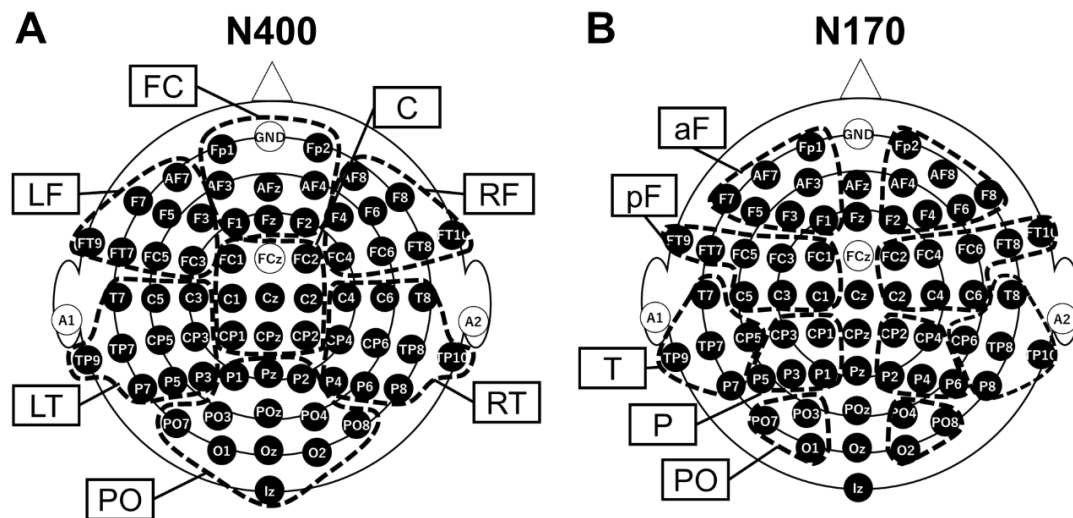


Figure 8. ROIs for N400 (A) and N170 (B)

A. ROIs for the N400.

FC, front-central; LF, left frontal; RF, right frontal; C, central; LT, left temporal; RT, right temporal; PO, parieto-occipital.

B. ROIs for the N170.

aF, anterior frontal; pF, posterior frontal; T, temporal; P, parietal; PO, parieto-occipital. We analyzed the corresponding regions on the right, while midline electrodes were excluded from the analyses.

For the analyses of the N170, epoched ERPs were averaged for the following five ROIs on the left following Lavric, Clapp, and Rastle (2007): anterior frontal (**aF**: Fp1, AF3, AF7, F1, F3, F5, F7), posterior frontal (**pF**: FC1, FC3, FC5, FT7, FT9, C1, C3, C5), temporal (**T**: T7, CP5, TP7, TP9, P7), parietal (**P**: CP1, CP3, P1, P3, P5) and parieto-occipital (**PO**: PO3, PO5, O1). We also analyzed the corresponding regions on the right, while midline electrodes were not analyzed (Figure 8B). Following the result of experiment 1 and other studies of the N170/M170 (Morris et al., 2013; Fruchter et al., 2013), the time window for the N170 was restricted to 150–200 ms. To

identify the generation source of the N170, we conducted a two-way rANOVA (**Laterality** [**Left** and **Right**] $\times$ **ROIs** [**aF**, **pF**, **T**, **P**, and **PO**]) for the average amplitude of the N170 time windows. When we found significant main effects or interactions between ROIs and laterality, we performed a follow-up ANOVA (**Prime Type** [**Related** and **Unrelated**] $\times$ **Suffix** [**-sa** and **-mi**]) to assess the presence of priming effects in each condition.

For the *-mi* type (*-sa* and *-mi* nouns), We conducted linear mixed-effects (LME) model analyses to examine the influence of individual target variables (i.e., TP, frequency, and length) on the moving average of a 5 ms change in millisecond-level amplitude over the time window of the N400. We used the `lmer` function in the `lme4` package (Bates et al., 2015) and the `clusterperm.lmer` function in the `permutes` package of R (Voeten, 2022). We used the maximal feasible models, that is, the maximal model that is still capable of converging (Bates et al., 2018; Voeten, 2022). Importantly, as the TP is always less than one, the log-transformed TP, which was used in linear mixed-effects modeling, is always a negative value (see 3.2.2. Stimuli for the definition of the TP). In addition, as the N400 is a negative-going potential, the positive regression coefficient (beta) in linear mixed-effects models for the TP indicates a larger N400, while negative beta indicates a smaller N400.

We conducted two mixed-effects model analyses of ROI amplitude to examine the influence of individual target variables (i.e., TP, frequency, and length) on the moving average of a 5 ms change in millisecond-level amplitude over the time window of the N400. We used the `lmer` function in the `lme4` package (Bates et al., 2015) and the `clusterperm.lmer` function in the `permutes` package of R (Voeten, 2022). We used two equations for the mixed-effects model analyses because of the high correlation between TP and frequency: one equation included TP and length as fixed effects (Equation (6)), and the other equation included frequency and length as fixed effects (Equation (7)):

$$Amplitude \sim TP + length + (1|Subject) \quad (6)$$

$$Amplitude \sim frequency + length + (1|Subject) \quad (7)$$

In addition, for the non-*mi* types (-*sa* nouns), to investigate the effect on morphological complexity for the N400, additional LME analysis is conducted for Equation (6) as the non-*mi* types are always -*sa* nouns, which both theories predict decomposition for them.

Considering the numerous regressions computed in the analysis, we performed a multiple comparison correction procedure applied in experiment 1 (3.2.6.). In addition, we computed the beta values of fixed effects for 10,000 random permutations to identify temporal clusters in

which the target variables were significantly correlated with the amplitude. To investigate the effect of priming on the ERPs, we conducted this analysis separately for Prime type (i.e., **Related** and **Unrelated** conditions).

4.3 Results

In section 4.3.1, we conduct behavioral and ERP analyses for the *-mi* type to examine whether the lexicalized *-mi* noun dealt with their morphemes or whole word in the stage of lexical processing (4.3.1). Specifically, we conduct analyses for the ERPs of the two nouns. First, we conduct a three-way rANOVA (**Prime type** $\times$ **Suffix** $\times$ **ROI**) to examine the significance of the main effects or interactions. Then, we applied the LME model to estimate the complexity modulation for the N400. In addition, the analysis for N170 laterality was conducted.

In section 4.3.2, we conduct additional ERP (N400) analyses for non-*mi* types (*-sa* nouns), which differ in their complexities (e.g., “*naga-sa*” and “*hoso-naga-sa*,” length and slenderness (being long and thin)). Based on the analysis of the *-mi* type (4.3.1), we examine how the complexity of words affects the N400 compared to the non-*mi* type with a simple stem and non-

mi type with a complex stem. Since the non-*mi* type has two types of stems that differ in the number of morphemes, this section provides how the modulation on the N400 differs between TP and complexity.

4.3.1 Results 1: *-mi* type (*-sa* and *-mi* nouns)

The summary of the behavioral data is illustrated in Table 7. The mean accuracy of the lexical decision task was 91.5%, indicating that the participants correctly performed this task.

Table 7. Behavioral data and priming effects (*-mi* type)

The numbers on the top and bottom of each cell denote the mean value and standard error of the mean, respectively. The priming effect indicates the mean subtraction of reaction time (RT) of the unrelated condition minus the related condition.

	<i>-mi</i> type (<i>-sa</i> noun)		<i>-mi</i> type (<i>-mi</i> noun)		Non <i>mi</i> type with a simple stem (<i>-sa</i> nouns)		Non- <i>mi</i> type with a simple stem (<i>-sa</i> nouns)	
	RT (ms)	Accur acy (%)	RT (ms)	Accur acy (%)	RT (ms)	Accur acy (%)	RT (ms)	Accur acy (%)
Related	665 ±25	96.26 ±0.9	735 ±32	86.10 ±2.0	95.3 ± 1.4	674 ± 25	86.8 ± 2.8	788 ± 32
Unrelated	757 ±31	95.49 ±1.3	793 ±34	84.20 ±2.8	94.0 ± 1.0	748 ± 27	81.4 ± 3.0	846 ± 32
Priming effect	−92	1.27	−58	1.90	1.3	−74	5.4	−58

4.3.1.1 Behavioral results 1: *-mi* type (*-sa* and *-mi* nouns)

Regarding the accuracy, a two-way rANOVA (**Prime type** $\times$ **Suffix**), the main effect of Suffix was significant for the analysis of the *-mi* type [**Suffix**: $F(1, 17) = 29.06, p < .0001, \eta^2_G = .31$] (Table 8). Conversely, the main effect of Prime type and interaction was not significant, suggesting that the participants had more difficulty with *-mi* nouns in the lexical decision task [**Prime type**: $F(1, 17) = 2.457, p = .14, \eta^2_G = .01$; **interaction**: $F(1, 17) = .07, p = .80, \eta^2_G = .0004$].

Table 8. ANOVA results for accuracy
df, degrees of freedom; η^2_G , generalized eta squared.

Factors	<i>df</i>	<i>F</i>	<i>p</i>	η^2_G
Prime type $\times$ Suffix				
Prime type	1, 17	2.46	.14	.010
Suffix	1, 17	29.06	<.0001	.31
Interaction	1, 17	0.069	.80	.0004
Prime type $\times$ Stem type				
Prime type	1, 17	11.87	.0031	.027
Stem type	1.21, 20.6	17.74	.0002	.31
Interaction	1.77, 30.03	2.92	.076	.014

An additional two-way rANOVA (**Prime Type** $\times$ **Stem Type**) on the accuracy demonstrated the significant main effects of prime type and stem type [**Prime Type**: $F(1, 17) = 11.87, p = .0031, \eta^2_G = .027$; **Stem Type**:

$F(1.21, 20.6) = 17.74, p = .0002, \eta^2_G = .31$], while the interaction between prime type and stem type was marginally significant [**interaction:** $F(1.77, 30.03) = 2.92, p = .076, \eta^2_G = .014$] (Table 8). Multiple comparisons for the stem types demonstrated that the non-*mi* type with a complex stem had a lower accuracy than other stem types [**-*mi* type vs. non-*mi* type with a complex stem:** $t(17) = 4.26, \text{corrected } p = .0009, \text{Cohen's } d = 1.24$; **Non-*mi* type with a simple stem vs. non-*mi* type with a complex stem:** $t(17) = 4.53, \text{corrected } p = .0009, \text{Cohen's } d = 1.12$], whereas the -*mi* type and non-*mi* type with a simple stem had no significant difference in accuracy [$t(17) = 1.02, p = .32, \text{Cohen's } d = 0.22$].

Regardless of the target conditions, the reaction times were significantly facilitated when the morphologically related prime preceded the target (Table 7). For a two-way ANOVA (**Prime type** $\times$ **Suffix**) of the -*mi* type, we found facilitating priming effects for the related conditions [**Prime type:** $F(1, 17) = 70.11, p < .0001, \eta^2_G = .080$] (Table 9). Additionally, we found the significant main effect of suffix [**Suffix:** $F(1, 17) = 30.92, p < .0001, \eta^2_G < .001$]. Post-hoc analysis for Suffix demonstrated that the reaction times of -*mi* nouns were significantly longer than those of -*sa* nouns [**Related:** $F(1, 17) = 20.93, p = .0003, \eta^2_G = .0037$; **Unrelated:** $F(1, 17) = 6.696, p = .019, \eta^2_G = .0037$]. In contrast, the interaction between the Prime

types and Suffix was not significant [**interaction**: $F(1, 17) = 2.23$, $p = .15$, $\eta^2_G = .0037$]. These results demonstrated that both types of nouns showed a morphological priming effect, and the effects between the two types of nouns were equivalent.

Table 9. ANOVA results reaction time (-mi type)
 df , degrees of freedom; η^2_G , generalized eta squared.

Factors	df	F	p	η^2_G
Prime type $\times$ Suffix				
Prime type	1, 17	70.11	<.0001	.080
Suffix	1, 17	30.92	<.0001	.041
Interaction	1, 17	2.23	.15	.0042
Prime type $\times$ Stem type				
Prime type	1, 17	160.2	< .0001	.090
Stem type	1.87, 31.8	105.0	< .0001	.15
Interaction	1.88, 32.03	1.91	.17	.0032

The post-hoc analysis of the prime types demonstrated facilitating priming effects for the primed conditions [**-sa nouns**: $F(1,17) = 59.78$, $p < .0001$, $\eta^2_G = .038$; **-mi nouns**: $F(1, 17) = 12.87$, $p = .0023$, $\eta^2_G = .0037$]. Additionally, the main effect of Suffix was significant [**Suffix**: $F(1, 17) = 30.92$, $p < .0001$, $\eta^2_G = .041$]. The post-hoc analysis for the suffix demonstrated that the reaction times of -mi nouns were significantly longer than those of -sa nouns [**Related**: $F(1, 17) = 20.93$, $p = .0003$, $\eta^2_G = .0037$; **Unrelated**: $F(1,17) = 6.696$, $p = .019$, $\eta^2_G = .0037$]. In contrast, the interaction between the Prime types and Suffix was not significant [**interaction**: $F(1, 17) = 2.228$, $p = .15$, $\eta^2_G = .0037$]. These results

demonstrated that both types of nouns demonstrated a morphological priming effect, and the effects between the two types of nouns were equivalent.

4.3.1.2 ERP results 1: *-mi* type (*-sa* and *-mi* nouns)

4.3.1.2.1 Priming effects on the N400

The ERPs of the *-mi* type revealed attenuation of the N400 for the related condition compared to the unrelated condition in the C, LT, RT, and PO regions by visual inspection (Figure 9). The scalp distribution of the related-unrelated condition demonstrated that the unprimed condition was more negative in both types of nouns (Figure 10). A three-way rANOVA (**Prime type** $\times$ **Suffix** $\times$ **ROI**) for the *-mi* type of the mean amplitude of the N400 revealed a significant main effect of Prime type, Suffix and ROI [**Prime type**: $F(1, 17) = 20.96, p = .0003, \eta^2_G = .045$; **Suffix**: $F(1, 17) = 6.73, p = .019, \eta^2_G = .011$; **ROI**: $F(1.74, 29.59) = 8.59, p = .0017, \eta^2_G = .13$] (Table 10). Conversely, the main effect of the suffix was not significant [**Suffix**: $F(1, 17) = 2.851, p = .1096, \eta^2_G = .0113$]. The interaction between Prime type and ROI was also significant [**Prime type** $\times$ **ROI**: $F(2.48, 42.24) = 28.28, p$

$< .0001, \eta^2_G = .025$]. However, the interactions were not significant [**Prime type** $\times$ **Suffix**: $F(1, 17) = 0.1123, p = .74, \eta^2_G = .0002$, **Suffix** $\times$ **ROI**: $F(1.91, 32.54) = 1.170, p = .32, \eta^2_G = .0013$, **Prime type** $\times$ **Suffix** $\times$ **ROI**: $F(2.92, 49.68) = 1.905, p = .14, \eta^2_G = .0010$].

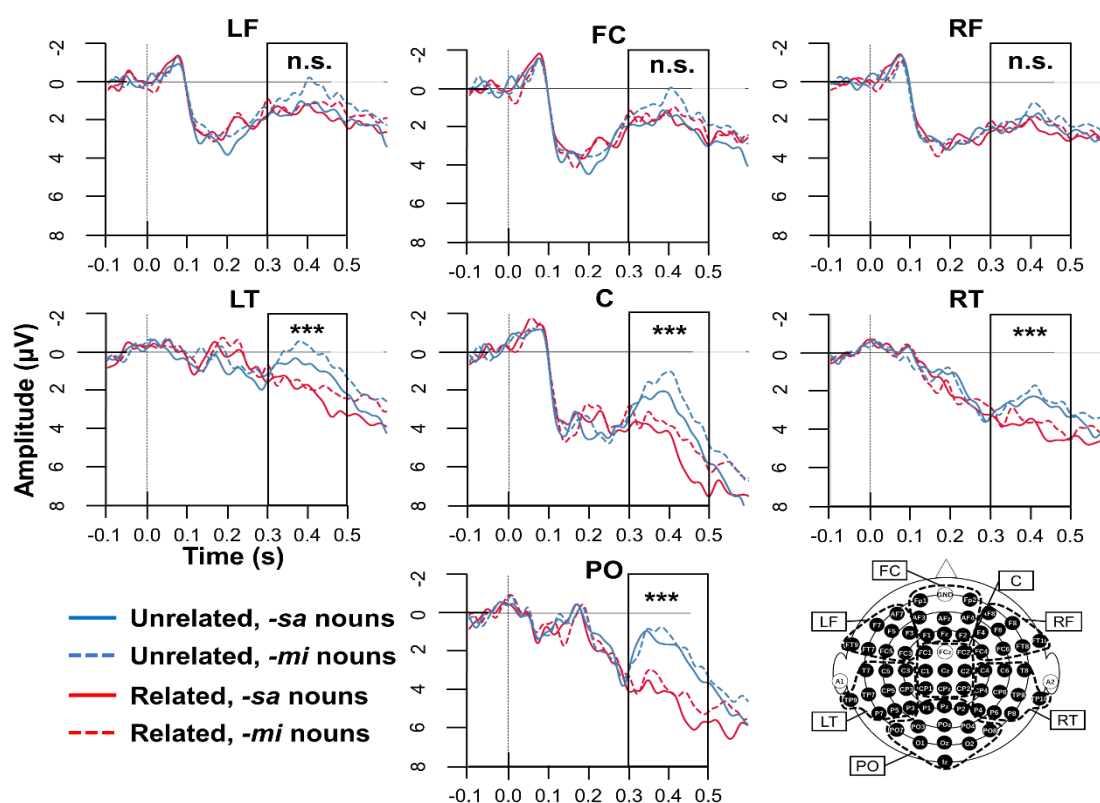


Figure 9. Grand-averaged ERPs of the four factors of the *-mi* type. Red and blue lines represent ERPs of related and unrelated conditions, respectively. Solid and broken lines represent ERPs of *-sa* nouns and *-mi* nouns, respectively. The rectangle areas (300–500 ms) indicate the time window of the N400. FC, fronto-central; C, central, LF, left frontal; RF, right frontal; LT, left temporal; RT, right temporal; PO, parieto-occipital. We applied the Greenhouse-Geisser correction for rANOVA and Shaffer’s modified sequentially rejective Bonferroni procedure for the post-hoc tests. *corrected $p < .05$, **corrected $p < .01$, ***corrected $p < .001$.

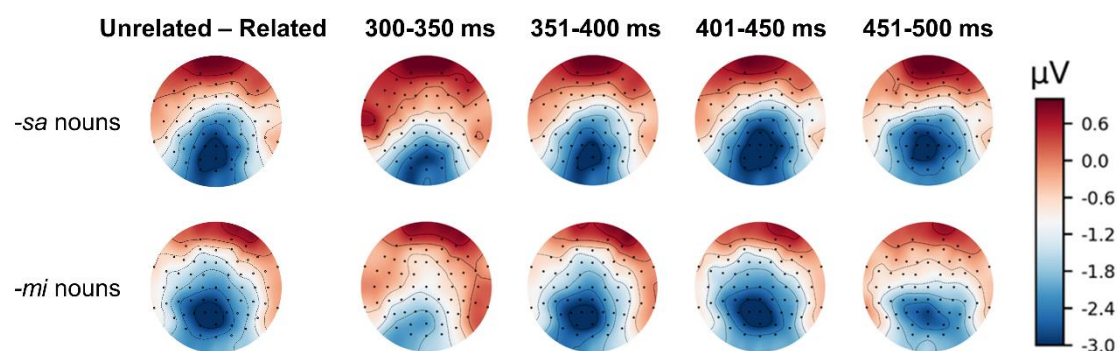


Figure 10. The scalp distribution of the unrelated – related

Table 10. ANOVA results of the N400 (-mi type)
df, degrees of freedom; η^2_G , generalized eta square

Factors	<i>df</i>	<i>F</i>	<i>p</i>	η^2_G
Prime Type $\times$ Suffix $\times$ ROI				
Prime Type	1, 17	20.96	.0003	.045
Suffix	1, 17	6.73	.019	.011
ROI	1.74, 29.59	8.59	.0017	.13
Prime Type $\times$ Suffix	1, 17	0.11	.74	.0002
Prime Type $\times$ ROI	2.48, 42.24	28.28	<.0001	.025
Suffix $\times$ ROI	1.91, 32.54	1.17	.32	.0013
Prime Type $\times$ Suffix $\times$ ROI	2.92, 49.68	1.91	.14	.0010
C: Prime Type $\times$ Suffix				
Prime Type	1, 17	22.10	.0002	.059
Suffix	1, 17	24.27	.015	.015
Prime Type $\times$ Suffix	1, 17	0.0066	.94	<.0001
LT: Prime Type $\times$ Suffix				
Prime Type	1, 17	24.27	.0001	.097
Suffix	1, 17	5.021	.039	.017
Prime Type $\times$ Suffix	1, 17	0.48	.50	.0013
RT: Prime Type $\times$ Suffix				
Prime Type	1, 17	23.69	.0001	.066
Suffix	1, 17	1.26	.28	.0043
Prime Type $\times$ Suffix	1, 17	0.20	.66	.0003
PO: Prime Type $\times$ Suffix				
Prime Type	1, 17	61.41	<.0001	.12
Suffix	1, 17	1.81	.20	.0044
Prime Type $\times$ Suffix	1, 17	0.52	.48	.0006

To determine conditions for further analyses, simple effects tests for significant interaction, i.e., Prime type and Suffix, were conducted. The effect of ROI was significant regardless of Prime type [**Related:** $F(1.87, 31.82) = 13.66$, $p = .0001$, $\eta^2_G = .19$, **Unrelated:** $F(1.7, 28.88) = 5.064$, p

= .017, $\eta^2_G = .096$]. Prime type was significant only in the C, LT, RT, and PO [C: $F(1, 17) = 22.10, p = .0002, \eta^2_G = .059$, LT: $F(1, 17) = 24.27, p = .0001, \eta^2_G = .097$, RT: $F(1, 17) = 23.69, p = .0001, \eta^2_G = .066$, PO: $F(1, 17) = 61.41, p < .0001, \eta^2_G = .12$] (Figure 11). While it was not significant in the FC, LF, and RF [FC: $F(1, 17) = 0.73, p = .40, \eta^2_G = .0031$, LF: $F(1, 17) = 2.851, p = .11, \eta^2_G = .0010$, RF: $F(1, 17) = 1.168, p = .29, \eta^2_G = .0050$]. These results suggest that the N400, hallmark of lexical processing, is only restricted in the C, LT, RT, and PO.

We conducted follow-up two-way rANOVAs (**Prime Type** $\times$ **Suffix**) for each of the C, LT, RT, and PO regions, demonstrating significant priming effects. In addition to the significant priming effects, we found a significant main effect of Suffix in the C and LT [C: $F(1, 17) = 7.30, p = .015, \eta^2_G = .015$; LT: $F(1, 17) = 5.02, p = .039, \eta^2_G = .017$] (Figure 11 and Table 10). However, the RT and PO regions did not demonstrate a significant effect of Suffix [RT: $F(1, 17) = 1.26, p = .28, \eta^2_G = .0043$; PO: $F(1, 17) = 1.81, p = .20, \eta^2_G = .0044$]. These results suggest that lexical (semantic) processing related to the lexicalized meaning of *-mi* nouns elicited the larger N400 in the C and LT regions. The interaction was not significant in all regions [C: $F(1, 17) =$

0.0066, $p = .94$, $\eta^2_G < .0001$; **LT**: $F(1, 17) = 0.48$, $p = .50$, $\eta^2_G = .0013$; **RT**:
 $F(1, 17) = 0.20$, $p = .66$, $\eta^2_G = .0003$; **PO**: $F(1, 17) = 0.52$, $p = .48$, η^2_G
 $= .0006$].

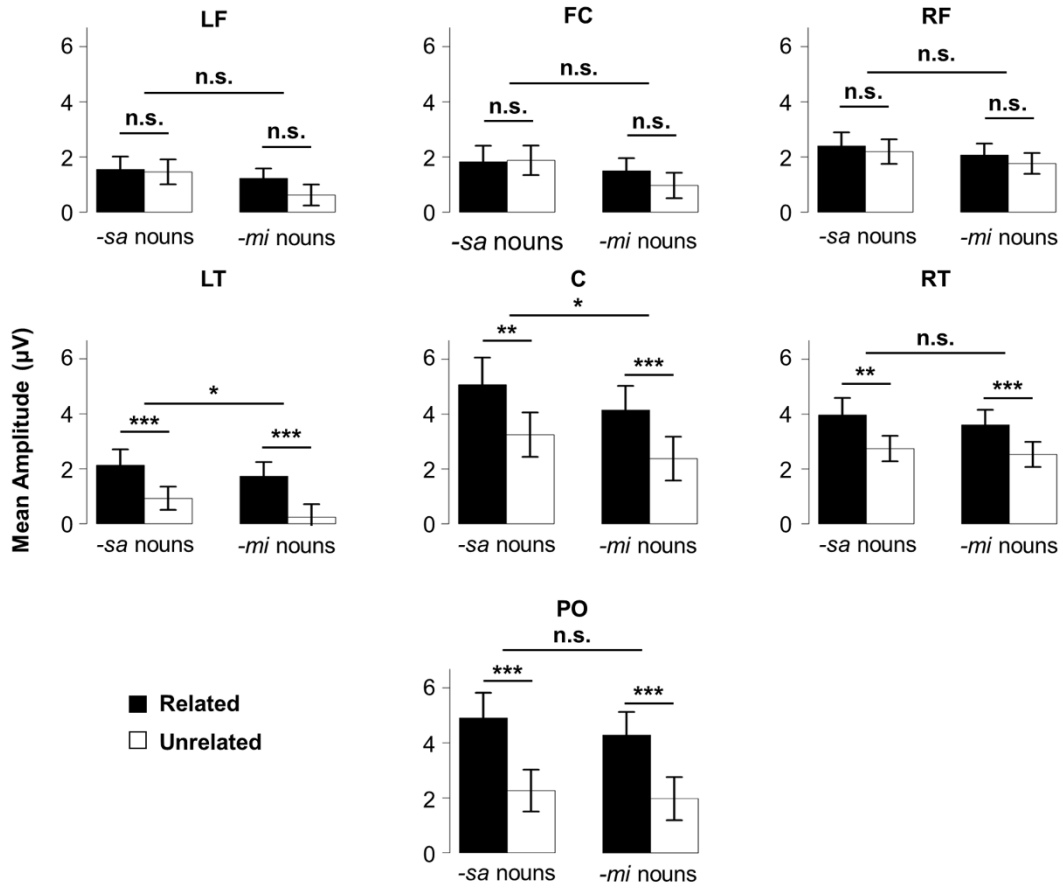


Figure 11. Mean amplitude of N400 between conditions for each ROIs (Prime type: black and white, Suffix: right and left)
The amplitudes of the N400 (mean $\pm$ SEM). Filled and open bars denote the related and unrelated conditions, respectively. We applied the Greenhouse-Geisser correction for rANOVA and Shaffer's modified sequentially rejective Bonferroni procedure for the post-hoc tests. *corrected $p < .05$, **corrected $p < .01$, ***corrected $p < .001$.

4.3.1.2.2 Attenuation of the N400 by TP

The linear mixed-effects model analyses revealed that the model including the TP between the stem and suffix as a fixed effect demonstrated significant negative correlations for the unrelated conditions in two different time windows (Figure 12A). The correlation between the N400 amplitude and TP was significant in the C, LT, and PO regions around 340–350 ms, while around 440–460 ms, the correlation was significant in the C, LT, RT, and PO regions (corrected $p < 0.05$). As we found the negative regression coefficient (beta) in these regions, these results demonstrated that the TP between morphemes attenuated the amplitude of the N400, which is a crucial ERP component for lexical access to morphemes and recombination of morphemes (see 4.2.6. Data processing for the relationship between the N400 and regression coefficient of the TP). In contrast, no significant correlations were found for the related condition (Figure 12B). We further tested another model including frequency as a fixed effect, which did not show a significant correlation in any regions or conditions throughout the N400 time window (300–500 ms).

To determine the conditions for further analyses, simple effects tests for the significant interactions—Prime type and Suffix—were conducted. ROI was significant regardless Prime type [**Related:** $F(1.87, 31.82) = 13.66$, p

= .0001, $\eta^2_G = .1908$, **Unrelated**: $F(1.7, 28.88) = 5.064$, $p = .0168$, $\eta^2_G = .0960$]. Prime type was significant only in the C, LT, RT, and PO [C: $F(1, 17) = 22.10$, $p = .0002$, $\eta^2_G = .0587$, **LT**: $F(1, 17) = 24.27$, $p = .0001$, $\eta^2_G = .0966$, **RT**: $F(1, 17) = 23.69$, $p = .0001$, $\eta^2_G = .0659$, **PO**: $F(1, 17) = 61.41$, $p < .0001$, $\eta^2_G = .1160$] (Table 10). Prime type was not significant in the CF, LF, and RF [**FC**: $F(1, 17) = 0.7341$, $p = .4035$, $\eta^2_G = .0031$, **LF**: $F(1, 17) = 2.851$, $p = .1096$, $\eta^2_G = .0099$, **RF**: $F(1, 17) = 1.168$, $p = .2948$, $\eta^2_G = .0050$].

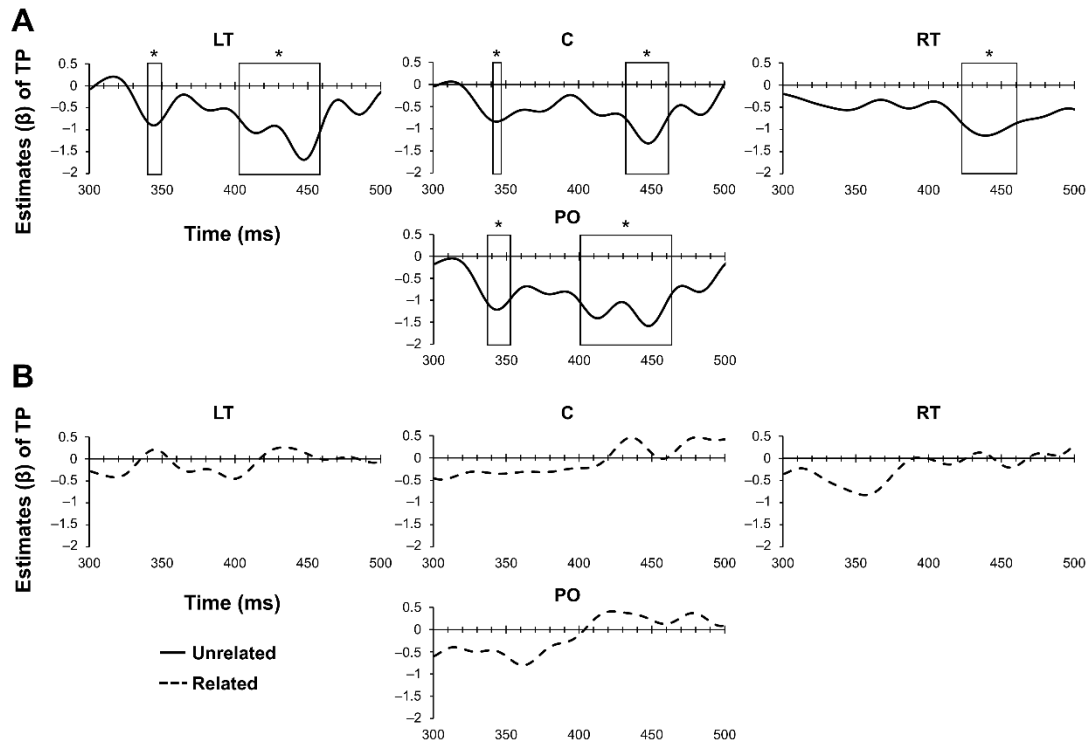


Figure 12. Effect of the TP between Stem and Suffix in the N400 regions
A: Related conditions. **B:** Unrelated conditions.

The rectangle areas indicate significant temporal clusters (corrected $p < 0.05$). C, central; LT, left temporal; RT, right temporal; PO, parieto-occipital. We performed a multiple comparison correction procedure on the continuous moving average amplitude of the point-by-point regressions using a permutation test. *corrected $p < .05$.

4.3.1.2.3 Significant laterality effects on the N170

Among all four conditions, the scalp distribution of the N170 had a lower amplitude in the left hemisphere than in the right hemisphere, indicating that the N170 component was localized in the left hemisphere (Figure 13). A two-way rANOVA (**Laterality** $\times$ **ROIs**) revealed significant main effects for laterality and ROIs [**Laterality**: $F(1, 71) = 7.37, p = .0083, \eta^2_G = .0081$; **ROIs**: $F(1.76, 125.0) = 24.77, p < .0001, \eta^2_G = .084$] (Table 11). Additionally, the interaction between laterality and ROI was also significant [**interaction**: $F(2.42, 172.2) = 6.149, p = .0013, \eta^2_G = .0025$]. Multiple comparisons among the ROIs demonstrated significantly lower amplitudes for the T and PO regions than for other regions [**T vs. aF**: $t(71) = 7.16$, corrected $p < .0001$, Cohen's $d = 0.81$; **T vs. pF**: $t(71) = 8.48$, corrected $p < .0001$, Cohen's $d = 0.70$; **T vs. P**: $t(71) = 4.09$, corrected $p = .0005$, Cohen's $d = 0.34$; **PO vs. aF**: $t(71) = 5.06$, corrected $p < .0001$, Cohen's $d = 0.72$; **PO vs. pF**: $t(71) = 5.94$, corrected $p < .0001$, Cohen's $d = 0.65$; **PO vs. P**: $t(71) = 5.99$, corrected $p < .0001$, Cohen's $d = 0.37$]. In contrast, the post-hoc analysis revealed no significant difference between the mean amplitudes of T and PO [$t(71) = 1.19$, corrected $p = .48$, Cohen's $d = 0.10$]. Comparing the ERPs of these regions, a clear sharp peak was found in the PO region of the left hemisphere (Figure 13). We found significant laterality effects on the

N170 in the T and PO regions [**T**: $t(71) = 2.56$, corrected $p = .013$, Cohen's $d = 0.30$; **PO**: $t(71) = 3.32$, corrected $p = .0028$, Cohen's $d = 0.39$].

A follow-up rANOVA (**Prime Type** $\times$ **Suffix**) was performed for PO in the left hemisphere. No significant main effects of Prime type and Suffix or interaction were found [**Prime Type**: $F(1, 17) = 0.038$, $p = .85$, $\eta^2_G = .0006$; **Suffix**: $F(1, 17) = 0.011$, $p = .92$, $\eta^2_G = .0002$; **interaction**: $F(1, 17) = 0.079$, $p = .78$, $\eta^2_G = .0011$] (Table 11). These results indicated that the target *-sa* nouns and *-mi* nouns were processed similarly in this time window, even if morphologically related primes preceded the nouns. In addition, the linear mixed-effects model analyses of the left PO amplitude found no significant temporal cluster regardless of the TP or Frequency (corrected $p > .05$), indicating that TP and frequency did not modulate the amplitude of the N170.

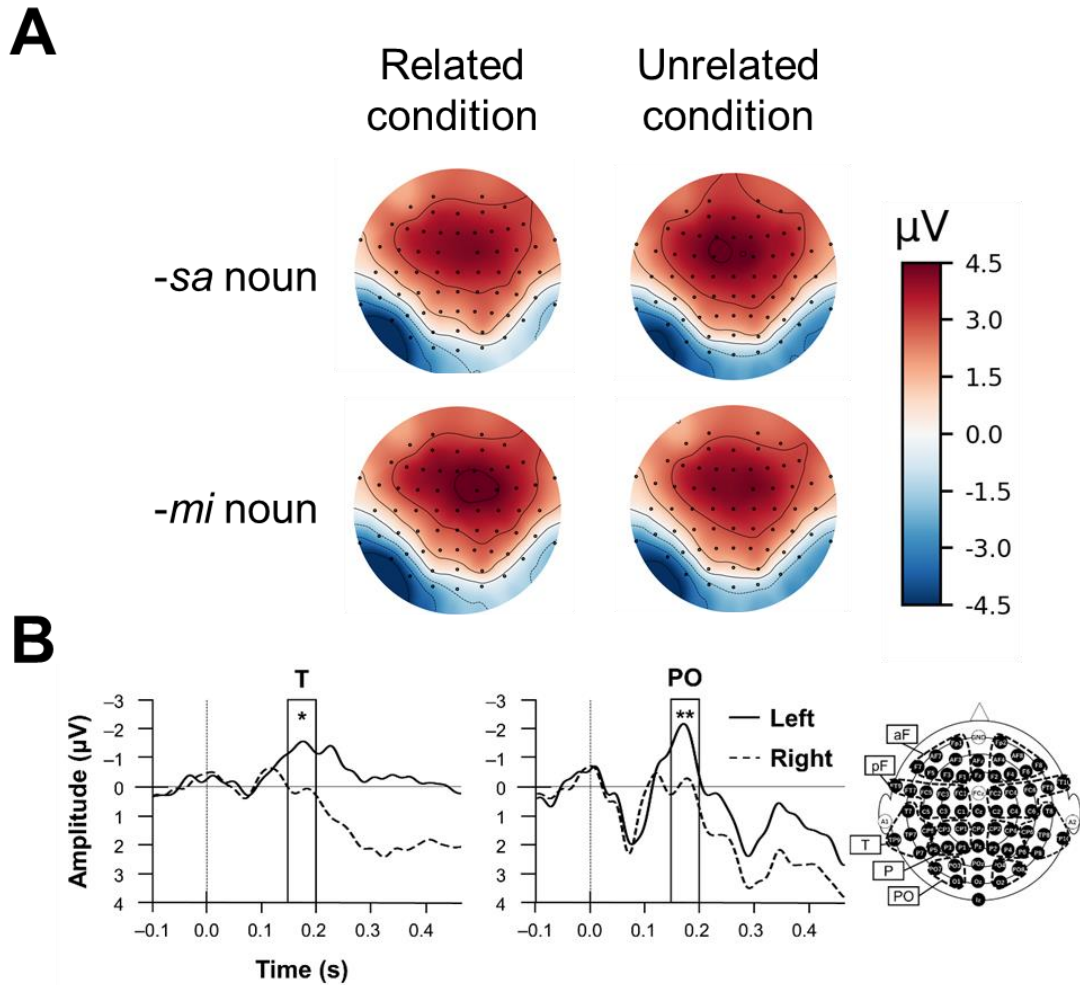


Figure 13. Significant laterality effects on the N170 in the temporal and parieto-occipital regions

A: Averaged N170 (150–200 ms) scalp distributions of the four factors (Prime type $\times$ Suffix). The N170 was more negative-going in the left T and PO regions.

B: ERPs of the PO and T regions. Solid and dashed lines represent ERPs of the left and right hemispheres, respectively. The rectangle areas indicate the time window of the N170 (150–200 ms). We applied the Greenhouse-Geisser correction for rANOVA and Shaffer's modified sequentially rejective Bonferroni procedure for the post-hoc tests. *corrected $p < .05$, **corrected $p < .01$.

Table 11. Statistical results of averaged N170 amplitude for *-mi* type
 df , degrees of freedom; η^2_G , generalized eta squared.

Factors	df	F	p	η^2_G
Laterality $\times$ ROIs				
Laterality	1, 17	7.371	.0083	.0081
ROIs	1.76, 124.97	24.77	<.0001	.0840
Interaction	2.42, 172.15	6.149	.0013	.0025
PO: Prime type $\times$ Suffix				
Prime type	1, 17	0.038	.85	.0006
Suffix	1, 17	0.011	.92	.0002
Interaction	1, 173	0.079	.78	.0011

4.3.2 ERP Results 2: non-*mi* type (-*sa* nouns)

In this section, we conduct additional analyses for non-*mi* types (-*sa* nouns: e.g., “*naga-sa*” and “*hoso-naga-sa*”) on the N400 to investigate further the neural profile from the viewpoint of the complexity of words (i.e., No. of morphemes).

Based on the N400 analysis for -*mi* type (4.3.1.2.1), ROIs (i.e., C, LT, RT, and PO) were specified for further analyses.

4.3.2.1 Priming effects on the N400

A visual inspection demonstrated lower amplitude for the N400 time window and the N400 attenuation was seen in the ROIs (LT, C, RT, and PO) (Figure 14). Using a two-way ANOVA (**Prime type** $\times$ **ROI**), for the non-*mi* type with a simple stem, we found that significant main effect for the Prime type. In addition, the related condition showed the attenuated peak of the N400 for the non-*mi* type. In addition, the main effect of ROI was also significant (Table 12) [**Prime type**: $F(1, 17) = 21.19, p = .0003, \eta^2_G = .10$; **ROI**: $F(2.41, 40.99) = 11.16, p = .0001, \eta^2_G = .11$; **interaction**: $F(1.76, 29.86) = 3.568, p = .046, \eta^2_G = .0088$]. The non-*mi* type with complex stem

demonstrated a significant main effect of the Prime type and ROI, whereas the interaction was not significant (Table 12) [**Prime type**: $F(1, 17) = 21.64$, $p = .0002$, $\eta^2_G = .14$; **ROI**: $F(2.67, 45.44) = 7.531$, $p = .0005$, $\eta^2_G = .14$; **interaction**: $F(2.16, 37.2) = 2.109$, $p = .13$, $\eta^2_G = .018$]. In addition, post-hoc comparisons between the related and unrelated conditions in each ROI demonstrated significant priming effects in all regions (Table 12) [**Non-*mi* type with a simple stem**: **C**: $t(17) = 3.26$, corrected $p = .0046$, Cohen's $d = 0.77$; **LT**: $t(17) = 4.40$, corrected $p = .0011$, Cohen's $d = 1.04$; **RT**: $t(17) = 3.52$, corrected $p = .0052$, Cohen's $d = 0.83$; **PO**: $t(17) = 5.34$, corrected $p < .0001$, Cohen's $d = 1.34$; **Non-*mi* type with a complex stem**: **C**: $t(17) = 2.38$, corrected $p = .029$, Cohen's $d = 0.56$; **LT**: $t(17) = 3.56$, corrected $p = .0073$, Cohen's $d = 0.84$; **RT**: $t(17) = 3.34$, corrected $p = .0078$, Cohen's $d = 0.79$; **PO**: $t(17) = 3.78$, corrected $p = .0060$, Cohen's $d = 0.89$].

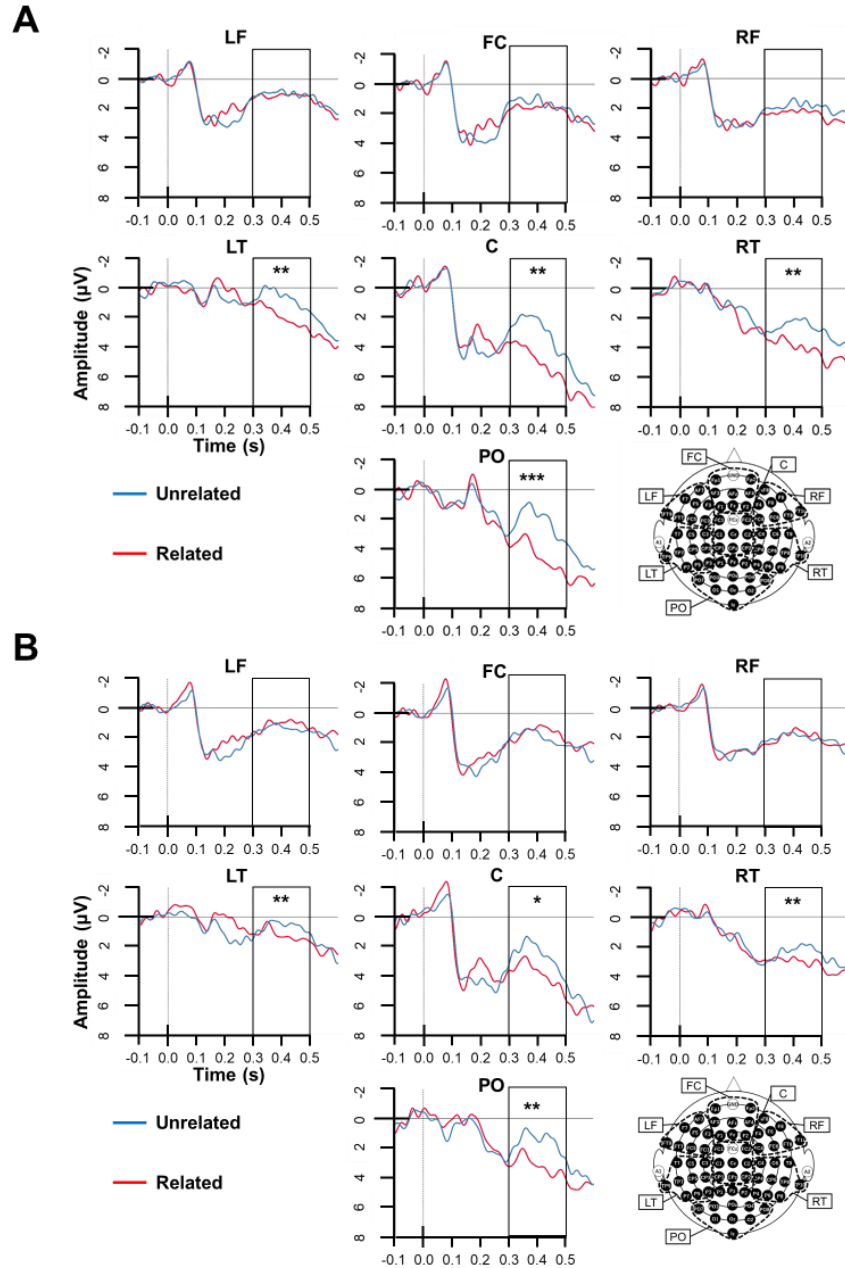


Figure 14. Grand-averaged ERPs of the four factors of *-mi* type
A: non-*mi* type with a simple stem, **B:** non-*mi* type with a complex stem
 Red and blue lines represent ERPs of related and unrelated conditions, respectively. The rectangle areas (300–500 ms) indicate the time window of the N400. FC, fronto-central; C, central, LF, left frontal; RF, right frontal; LT, left temporal; RT, right temporal; PO, parieto-occipital. We applied the Greenhouse-Geisser correction for rANOVA and Shaffer's modified sequentially rejective Bonferroni procedure for the post-hoc tests. *corrected $p < .05$, **corrected $p < .01$, ***corrected $p < .001$.

Table 12. ANOVA results of the N400 (non-*mi* type)*df*, degrees of freedom; η^2_G , generalized eta square

Factors	<i>df</i>	<i>F/t</i>	<i>p</i>	η^2_G /Cohen's <i>d</i>
simple stem				
Prime type × ROI				
Prime type	1, 17	21.19	.0003	0.10
ROI	2.41, 40.99	11.17	.0001	0.11
Interaction	1.76, 29.86	3.568	.461	0.0088
C	17	3.26	.0046	0.077
LT	17	4.40	.011	1.04
RT	17	3.52	.0052	0.083
PO	17	5.34	< .0001	1.34
complex stem				
Prime type × ROI				
Prime type	1, 17	21.6	.0002	0.14
ROI	2.67, 45.44	7.53	.0005	0.14
Interaction	2.19, 3.486	2.11	.13	0.018
C	17	2.38	.0029	0.56
LT	17	3.56	.0073	0.84
RT	17	3.34	.0078	0.079
PO	17	3.78	.0060	0.89

4.3.2.2 Attenuation of the N400 by TP

To investigate the complexity (**No. of morphemes**) on the N400, a linear mixed (LME) model analysis that includes the number of morphemes was conducted for non-*mi* stems. The equation of LME (8) is as follows:

$$Amplitude \sim No. of morpheme + TP + length + (1|Subject) \quad (8)$$

The LME model analyses revealed that the No. of morphemes as a fixed effect in Equation (8) demonstrated a significant positive correlation for the unrelated conditions around 350–360 ms in LT, C, and PO (Figure 15A, corrected $p < 0.05$). In addition, LT demonstrated another significant correlation around 430–450 ms. Interestingly, the time windows of significant correlations replicated the result of the *-mi* type (Figure 12A), which found two significant time windows, suggesting the No. of morphemes is also involved in the lexical processing (i.e., lexical access and recombination). As we found the positive regression coefficient (beta) in these regions, these results demonstrated that the larger number of morphemes makes the peak of the N400 smaller, which is a crucial ERP component for lexical access to morphemes and the recombination of morphemes. For the related condition, the No. of morphemes demonstrated

a correlation around 440–480 ms (Figure 15B). This TP result is inconsistent with the TP's effect for the *-mi* type. Additionally, the PO demonstrated another significant correlation in around 300–325 ms. As well as the related condition, the time window of beta in these regions coincided with TP's effect for the *-mi* type. Contrary to the related conditions, the unrelated condition did not correlate in any region's time window around 350 ms.

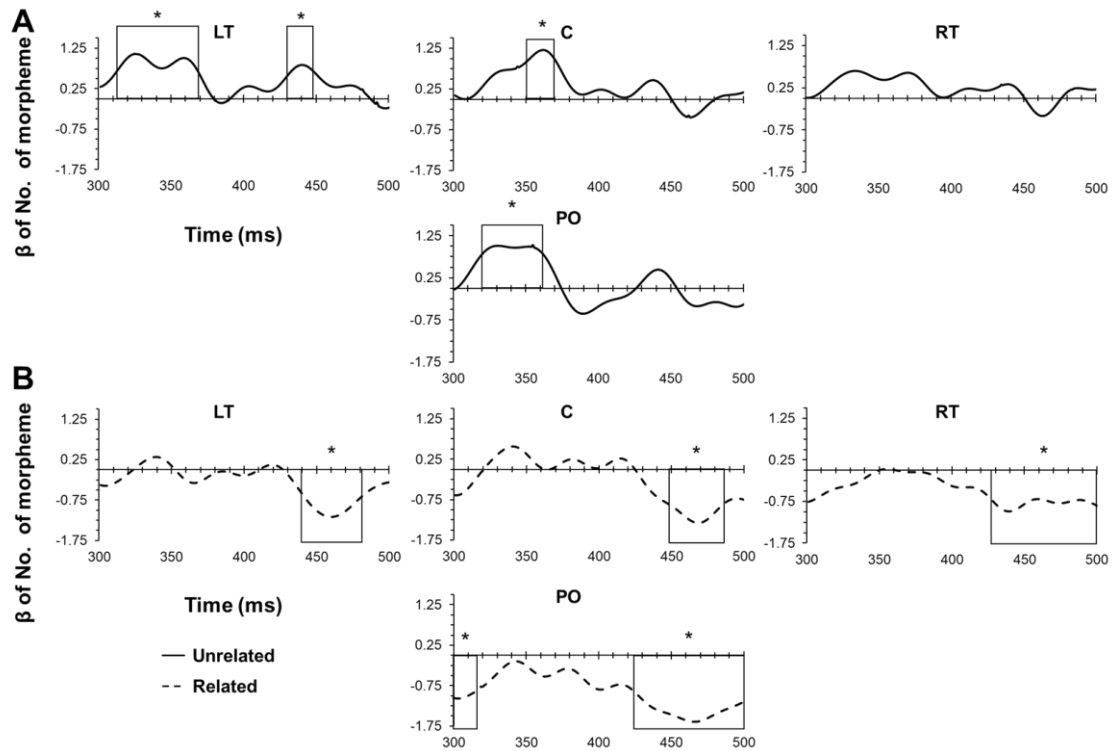


Figure 15. Effect of the No. of morpheme for non-*mi* type in the N400 regions

A: Related conditions. **B:** Unrelated conditions.

The rectangle areas indicate significant temporal clusters (corrected $p < 0.05$). C, central; LT, left temporal; RT, right temporal; PO, parieto-occipital. We performed a multiple comparison correction procedure on the continuous moving average amplitude of the point-by-point regressions using a permutation test. *corrected $p < .05$.

The LME model analyses revealed that the TP as a fixed effect demonstrated a significant positive correlation for the unrelated conditions around 440–450 ms in LT, C, and RT (Figure 16A, corrected $p < 0.05$). In addition, C, RT, and PO demonstrated another significant correlation around 330–350 ms. As well as the result of the No. of morphemes, the time windows of significant correlations replicated the result of the *-mi* type (Figure 12A). In contrast to the result of the *-mi* type, the regression coefficients (beta) in unrelated condition were positive in all ROIs (Figure 12A and Figure 15A). This result suggests that the TP affects oppositely to the stem type (i.e., *-mi* type or non-*mi* type). For the related condition, we found a significant correlation of around 320–380 ms for the related condition in LT, C, RT, and PO. Intriguingly, the correlation around 350 ms is lacking in the LT and C, which corresponds to the language area. This suggests that the priming lightens the correlation between N400 and the TP (i.e., the processing load) of around 350 ms in LT and C. In addition, It should be noted that the time window roughly corresponds to the significant correlations between Figure 12A and Figure 15A. For the related condition, the TP had a correlation of around 420–480 ms (Figure 16B). Besides,

contrary to the result of No. of morphemes, the beta/estimate values were positive, in the related condition. This demonstrates that the modulation of N400 between the number of morphemes and TP is different. Notably, the non-*mi* types showed significant correlations between the amplitude of N400 and fixed effects, even in unrelated conditions. Furthermore, the significant time windows for the unrelated conditions are distinct between the number of morphemes and TP (Figure 15B & Figure 16B).

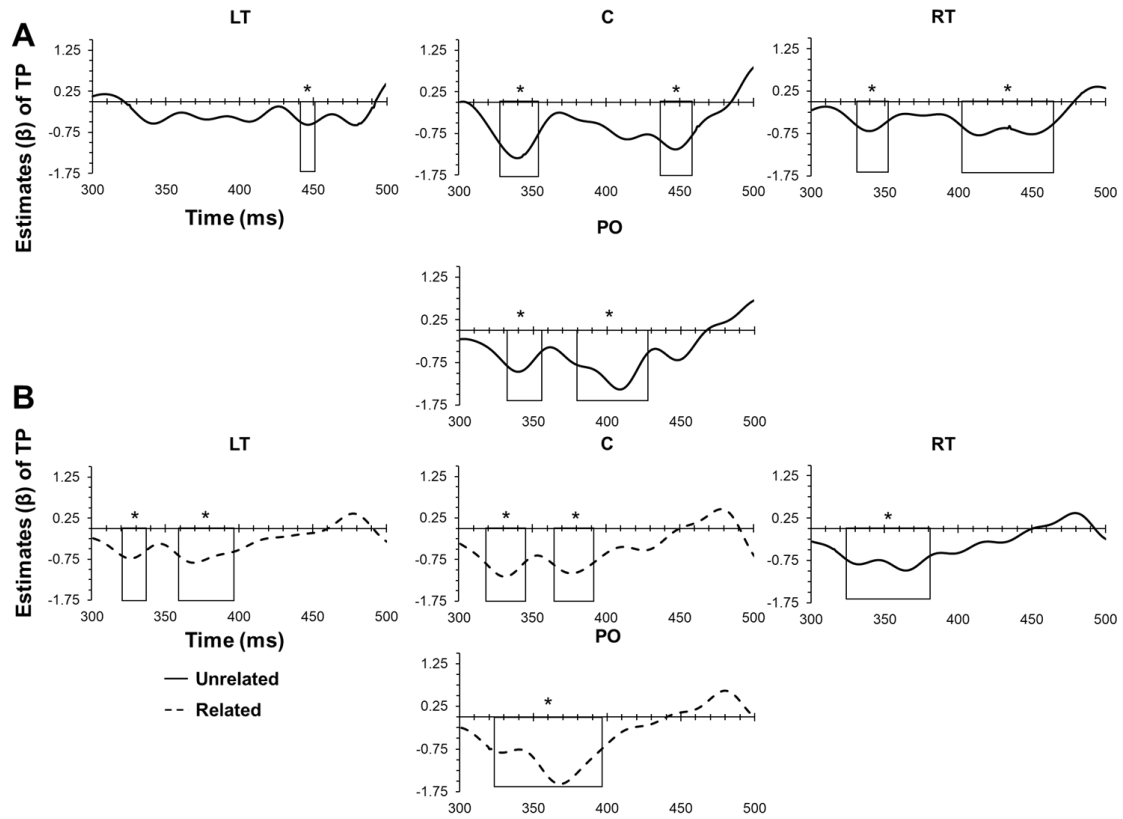


Figure 16. Effect of the TP for non-*mi* type in the N400 regions

A: Related conditions. **B:** Unrelated conditions.

The rectangle areas indicate significant temporal clusters (corrected $p < 0.05$). C, central; LT, left temporal; RT, right temporal; PO, parieto-occipital. We performed a multiple comparison correction procedure on the continuous moving average amplitude of the point-by-point regressions using a permutation test. *corrected $p < .05$.

4.4 Discussions

4.4.1 *-mi* type (*-sa* and *-mi* nouns)

In the present ERP experiment using the masked priming paradigm (Figure 7), we obtained three striking results from the analyses of the *-mi* type (4.3.1.). By examining the N400 and N170 components, we demonstrated the similarity and dissimilarity of two types of Japanese-derived nouns—semantically transparent *-sa* nouns and lexicalized *-mi* nouns (Table 13). First, we demonstrated the significant priming effects on the N400, as well as the significant laterality effects on the N170, for both *-sa* and *-mi* nouns (Figure 9, Figure 10, Figure 11, Figure 14, and Figure 12; Table 10, and Table 11). These results suggest that morphologically complex words were processed by a common neural mechanism, i.e., similarity. Moreover, we found a larger N400 and lower behavioral performance (lower accuracy and longer reaction times) for *-mi* nouns due to their lexicalized meaning, i.e., dissimilarity. Second, by using the linear mixed-effects models, we demonstrated that the TP from stem to suffix and the amplitude of the N400 had significant negative modulations in the LT, C, RT, and PO (corrected $p < .05$) (Figure 12). These results indicate that the TP attenuates

the amplitude of the N400. Third, we found that both types of derived nouns showed significant morphological priming effects in the behavioral data, that is, shorter reaction times and lower error rates under the related condition (Table 7, Table 8, and Table 9). These results also support similar neural mechanisms for processing these two types of derived nouns. In contrast, behavioral data also indicate that *-mi* nouns were more demanding than *-sa* nouns, further suggesting the dissimilarity of processing mechanisms for these de-adjectival nouns.

Table 13. Summary of similarities and dissimilarities between *-sa* and *-mi* nouns

	Similarities	Dissimilarities
Accuracy	—	<i>-sa</i> > <i>-mi</i>
Reaction times	Priming effects: Related < Unrelated	<i>-sa</i> < <i>-mi</i>
N400	Priming effects: Related < Unrelated	<i>-sa</i> < <i>-mi</i>
N170	Laterality effects: Left > Right	—

In the rANOVAs of the N400 for the factor of prime type, we found a significantly lower mean amplitude for the unrelated condition in the ROIs of C, LT, RT, and PO (Figure 9, Figure 10, and Table 10). These results demonstrated that the N400 was attenuated under the related condition, which was consistent with a previous masked priming ERP study (Lavric, at

al., 2007). As the masked prime activates the lexical entry of morphemes, these results also suggest the morphological relation between a prime adjective and a target *-sa/-mi* noun: that is, both the target *-sa* noun and *-mi* noun share a common adjectival stem with the prime adjective. As was reported in the previous ERP studies (Lavric et al., 2007; Morris & Stockall, 2012), these results further suggest that the common morpheme (i.e., stem) in the prime adjective and the target de-adjectival noun was activated in the parieto-occipital area. It should also be noted that the non-*mi* type stem showed similar attenuation of the N400 for related conditions (Figure 14). Intriguingly, we found a larger N400 for *-mi* nouns in the C and LT regions (Figure 11), indicating that the lexicalized meaning of *-mi* nouns elicited higher loads for semantic processing, which are consistent with previous studies (Hagiwara et al., 1999; Clahsen & Ikemoto, 2012). These results demonstrate the similarities and dissimilarities between *-sa* and *-mi* noun processing mechanisms.

In the mixed-effects model analyses of the N400, we found two significant temporal clusters in which the TP negatively modulated the N400 under the unrelated condition (Figure 12, and Figure 5). The surface frequency, however, did not have such a modulatory effect. This result implies that both types of derived nouns are dealt with as morphemes (stems

and suffixes), but not whole word forms. Further, the lack of temporal clusters in the related condition implies that the attenuated peak of the N400 shown in Figure 3A reflected the lower processing loads of morphological decomposition of the de-adjectival nouns. A previous study of the N400 component evoked by word processing suggests that there are some subcomponents in the N400 time window (Pykkänen & Marantz 2003). In addition, MEG studies that dealt with the M350 demonstrated that the M350 reflects two stages of the lexical process: lexical access and recombination of morphemes in this order (Solomyak & Marantz, 2009; Fruchter et al., 2013). Therefore, the two temporal clusters in the current experiment may correspond to two different stages of the lexical process. Importantly, linear mixed-effects modeling showed a negative beta value for the TP for each temporal cluster. In other words, since a larger N400 peak indicates a higher processing load for the target, this demonstrates that when the target TP is smaller, the following suffix becomes more difficult for readers to predict. These results demonstrate that derived nouns are represented as stems and suffixes, even for the *-mi* nouns, which contradicts Hagiwara et al. (1999) who postulated no decomposition.

In the behavioral data, we found robust priming effects in both types of nouns (Table 7, Table 8, and Table 9), which replicates the findings of

Clahsen and Ikemoto (2012). As semantic priming does not occur in the masked priming paradigm, the priming effect is understood to reflect the morpho-orthographical relation between the prime and target (Frost et al., 1997; Rastle et al., 2000). Intriguingly, the *-mi* nouns demonstrated longer reaction times and larger N400 (Figure 3C), reflecting lexical access and a morphological recombination, compared to the *-sa* nouns. These results demonstrate that a difference in semantic transparency can be observed at the behavioral and later neural level (i.e., >300 ms: the N400 time window in the experiment), but not at the earlier neural level (i.e., <200 ms: the N170 time window in the experiment). This interpretation can also explain the behavioral dissociation of *-sa* and *-mi* nouns for agrammatic aphasic patients (Hagiwara et al., 1999).

For the N170 component, which reflects morpho-orthographical decomposition, the rANOVA of ERPs revealed a sharp and significantly lower peak in the left PO region (Figure 13B). This result suggests that both noun types were decomposed into stems and suffixes. In addition, we found no significant difference between *-sa* nouns and *-mi* nouns, or between related and unrelated conditions (Figure 13A, and Table 11). These results further suggest that the target nouns are processed similarly, regardless of the noun type or prime type.

The current experiment demonstrated contradictory results to the previous study using cross-modal priming in aphasic studies, which support the dual mechanism theory (Hagiwara et al., 1999). Differences in the experimental paradigm may cause these inconsistent results. Namely, the previous studies used auditory and visual stimuli or required production and comprehension, while the present study uses visual stimuli and requires comprehension alone (Figure 6). In addition, we recruited a relatively small number of participants ($n = 18$), which could explain the lack of priming effect for the N170 (Figure 13A & Table 11), which was reported in the previous MEG study (Fruchter, et al., 2013). For further studies, we need to recruit more participants.

4.4.2 non-*mi* type (-*sa* nouns)

In the LME analyses of non-*mi* types, by examining the effect of the TP and the number of morphemes on the N400, we obtained four striking results from (4.3.2.). First, we demonstrated that the number of morphemes and amplitude of the N400 also demonstrated significant modulations in the LT, C, RT, and PO (corrected $p < .05$) (Figure 15). These results indicate that, as well as the TP, the number of morphemes also affects the lexical processing in the same time windows (i.e., lexical access and recombination). Second,

when the number of morphemes is included in the LME model, we found the correlations between the TP/number of morphemes and the N400 even in the related condition (Figure 12B, Figure 15B, & Figure 16B). This demonstrates that the attenuated peak of the N400 still reflects the TP and number of morphemes even though the related condition lightened the processing load. Third, consistent with the results above, we demonstrated that the correlation between the TP/number of morphemes and the N400 in the related condition is distinct through the time window of N400 in all ROIs. This result indicates that the TP and number of morphemes modulate the N400 in the different time stages. Fourth, by comparing the beta value of the related conditions, we found the difference in the modulation of the N400 between the number of morphemes and TP. The result demonstrates that the modulation on the N400 differs in the time stages and the beta value (negative and positive, respectively).

We found that, as well as the TP, the number of morphemes modulates the N400 (LT. C and PO). In addition, the modulation's time windows coincide with the TP of the *-mi* type (Figure 12A, Figure 15A, and Figure 16A). In the unrelated conditions, as we suggested above, the temporal clusters may correspond the lexical access and recombination (Solomyak & Marantz, 2009; Fruchter et al., 2013). This result which deals with another

aspect of complex words (i.e., complexity), further supports our suggestion that the correspondence between temporal clusters (correlations) and two stages of lexical processing. Moreover, it should be noted that only LT shows the latter temporal cluster, which possibly corresponds with the recombination (around 450 ms). This is consistent with MEG studies that have reported that the orbitofrontal cortex, which lies nearby the LT, activates with the process of recombination in this time window (Neophytou et al., 2018; Stockall et al., 2019). However, the positive beta value for the number of morphemes indicates that the larger the number of morphemes makes, the smaller peak in the N400. That is, the larger number of morphemes makes a smaller processing load in the lexical processing. This result is controversial and counterintuitive to the studies shown above, demonstrating larger activations for the compositional process. In this experiment, we used complex stem stimuli which differ in its structure (e.g., *[[hoso]-[naga]]-[sa]* and *[ko]-[[daka]-[sa]]*), we need to use controlled stimuli in its structure for further neurolinguistic studies of complexity (cf. Taft & Forster, 1975). As the studies in complex word processing have mainly used bimorphemic and monomorphemic words (Zweig & Pytkänen, 2009; Solomyak & Marantz, 2010; Lewis et al., 2011), these results, which deal with more complex trimorphemic words (e.g., “*hoso-naga-sa*”), need to

be further investigated by comparing the more broader number of morphemes (e.g., predict, predict-able, predict-abil-ity, and predict-abil-ities). In addition, the hierarchical structure of words (i.e., linear or nested; e.g., [[[predict]-abil]-ity] and [un-[[predict]-able]]) is also needed to be considered since some studies report the processing of prefixed/suffixed word difference in the brain activation (Zweig & Pylkkänen, 2009; Stockall et al., 2019).

In the LME analysis, we included the number of morphemes and TP as fixed effects (predictors) to investigate the effect of the number of morphemes on lexical processing. Contrary to the result of the N400, regardless of the fixed effects, we obtained that temporal clusters are found even in the Related conditions (Figure 12B vs. Figure 15B & Figure 16B). As reviewed in 4.1.2., the Related condition of masked priming lightens a load of the lexical process (Lavric et al., 2007; Fruchter et al., 2013), and the result of *-mi* type (Figure 12B) replicated the studies. However, the result of the non-*mi* types demonstrated temporal clusters in the related condition, which is inconsistent with the previous studies and the result of the current experiment. As mentioned above, as the number of morphemes has not been investigated in the previous neurolinguistic studies, this result suggests that involving the number of morphemes as a predictor provides a more plausible

model for morphologically complex word processing.

In the results of the Related condition, which lightens the load of lexical processing, we found that the temporal clusters were distinct between the fixed effect (i.e., number of morphemes and TP). Importantly, the TP correlates in the former temporal cluster around 350 ms, whereas the number of morphemes correlates with the latter temporal cluster around 450 ms (Figure 15B & Figure 16B). It is worth noting that the lack of correlation of a certain fixed effect in the Related condition indicates that the fixed effect is involved in the attenuation of the N400 in its temporal cluster. As a significant correlation explains the amplitude of the N400, the correlations of a certain fixed effect in the related condition reflect that the fixed effect explains the amplitude of the N400 in its temporal cluster. Therefore, the temporal clusters in the related condition in Figure 15B & Figure 16B indicate that the number of morphemes correlates with the former cluster.

In contrast, the TP correlates with the latter cluster. This replicates the studies that have reported that lexical access occurs in the early windows of the N400 around 300–400 ms (Pylkkänen & Marantz, 2003; Solomyak & Marantz, 2010). Intriguingly, the temporal cluster of the TP lacks around 350 ms in LT and C and the time window corresponds with the result of the *-mi* type (Figure 12A). In line with the discussion above, this result indicates that

the processing load of the TP in the N400 is lightened in this time window and in LT and C. As the corresponding component M350 is generated in the left temporal lobe, the lack of correlation on TP suggests that the left lateralization of the N400, which is observed in the centro-parietal area, can be observed by the pipeline of the current analysis. This result also supports the suggestion on the first result of the left lateralization of lexical processing in LT (Figure 15A).

By comparing the Unrelated and Related conditions, we found that the two different fixed effects modulate the N400 conversely. The beta of the number of morphemes for the N400 was positive and negative in the Unrelated and Related conditions. In contrast, the beta of TP for the N400 was positive in both conditions. This result demonstrates that the two fixed effects, the number of morphemes and TP, affect the N400 differently. To summarize, the difference between the number of morphemes and TP is observed in two ways, the temporal aspects, and the modulation.

4.5 Conclusions

4.5.1 *-mi* type (*-sa* and *-mi* nouns)

In the current experiment, we aim to fill the gap of inconsistent results in the two types of de-adjectival nouns, mainly focused on the N400 (4.3.1.)

Specifically, using the masked priming paradigm, we tried to elucidate whether the lexicalized *-mi* noun is treated as a genuinely dealt with the complex word (Introduction aim (1)). We hypothesize a common mechanism for two types of Japanese de-adjectival nouns and demonstrate this in an ERP study employing a masked priming paradigm. In the ERP analyses, the ERP components—N170 and N400—demonstrate that the neural processes between the two types of nouns are driven by a common mechanism (Figure 9, Figure 10, Figure 11, Figure 12, & Figure 13; Table 10 & Table 11). This is consistent with studies of complex word processing in English. In addition, we also found a larger N400 for *-mi* nouns due to their lexicalized meaning (Figure 11), suggesting the dissimilarity of the neural mechanism for processing two types of de-adjectival nouns. In particular, we elucidate where and when the variables of nouns modulate each ERP component by applying mixed-effects models (Figure 12). The modeling demonstrates two separate stages of the lexical process, which can be explained by the lemma transition probability (TP) between morphemes. In the behavioral analyses, we find a robust priming effect for the two types of nouns (Table 7, Table 8, & Table 9), supporting our hypothesis. Although further evidence is needed to establish a more detailed mechanism for word recognition, our neural and behavioral results clearly showed a common processing mechanism of the

two types of derived nouns in the earlier stage, i.e., morphological decomposition (Table 13). Whereas in the later stage, which reflects lexical access and recombination of morphemes, different neural mechanisms processed the two types of de-adjectival nouns (Table 13). These results demonstrate the similarities and dissimilarities of the neural mechanisms, which were also proposed in the past-tense debate in English, for processing two types of de-adjectival nouns in Japanese. By demonstrating that the Japanese word processing mechanism is consistent with that of typologically different languages, our study elucidates the cross-linguistic universality of the neural basis of morphological processing.

4.5.2 non-*mi* type (-*sa* nouns)

In the additional LME analyses, we further elucidated the neural profile of morphologically complex word processing from the viewpoint of the number of morphemes (i.e., complexity) (4.3.2.). The LME analyses revealed that the number of morphemes is an important indicator of the N400 and is clearly distinct from the TP in the time window and modulation of the lexical processing. In the LME analysis for the number of morphemes, we found a significant correlation between the number of morphemes and the N400, as well as the TP (Figure 15). This is consistent with the result of the

-mi type (Figure 12). In addition, we found a correlation between the fixed effects (i.e., number of morphemes and TP) and the N400 in the Related condition (Figure 15B & Figure 16B), suggesting that the number of morphemes is a crucial predictor of the N400. Furthermore, the temporal clusters were distinct between the fixed effect (i.e., number of morphemes and TP), indicating that the two fixed effects modulate the N400 differently in the time stages. The result suggests that the number of morphemes modulates the lexical access, and the TP modulates the recombination. Compared to the beta/estimate of the fixed effects, we found that the fixed effects are also different in the beta value. This provides another evidence of the difference between the two fixed effects. These results demonstrate the importance of the number of morphemes (complexity) in the lexical processing and the difference in the N400 modulation between the number of morphemes and TP. In particular, these results provide novel implications for processing complex words since the complexity has yet to be well investigated in previous neurolinguistic studies.

5. Experiment 3: Cortical Entrainment

5.1 Background

How language is processed in the brain is one of the fundamental problems for both neuroscience and linguistics. Regarding the relationship between theoretical linguistics and experimental neuroscience, previous studies pointed out that the linguistic theories must be coordinated to investigate its neural functions, further proposing “mapping problem,” i.e., the syntactic operations cannot be “mapped” in neural activations (Embick & Poeppel, 2015; Poeppel, 2012; Poeppel & Embick, 2005). However, recent neurophysiological studies suggested that cortical oscillations, i.e., rhythmic waves of brain activity, offer the solution to the mapping problem, as they track linguistically meaningful acoustic properties of speech (Giraud & Poeppel, 2012). Cortical entrainment, synchronization of neural oscillations with exogenous and isochronous rhythmic stimuli, was interpreted as information processing of external stimuli between brain areas (Buzsáki & Draguhn, 2004). A neurolinguistic study proposed that the different frequency bands of the cortical entrainment have different subfunctions of cortical information processing (e.g., processing of phonemes, syllables, phrases, and sentences) (Meyer, 2018). Martin and

Doumas (2017) and Martin (2020) further suggested that cortical entrainments induced by auditory exogenous linguistic stimuli could reflect the syntactic operations to the abstract hierarchical structure.

In auditory processing for phrases and sentences, Ding et al. (2016) demonstrated that the rhythmic presentation of words induced cortical entrainment in the delta (1–4 Hz) band frequency that corresponded to the hierarchical structures of words, phrases, and sentences. Moreover, they also reported that such entrainments occurred cross-linguistically (Mandarin Chinese and English) for native speakers but not for non-native speakers, who did not have linguistic knowledge of the target language. In addition, Lo et al. (2022) demonstrated that the entrainments reflected the hierarchy of these linguistic units, but not the patterns of words and word categories rhythmically repeated across the presented stimuli. Furthermore, these cortical entrainments were replicated in various neuroimaging studies using different methods, such as magnetencephalography (MEG) (Ding et al., 2016), electrocorticography (ECoG) (Ding et al., 2016), and electroencephalography (EEG) (Ding et al., 2017; Lo et al., 2022).

Meanwhile, whether the entrainment occurs in visual modality, which

is another important modality for language, is yet to be investigated because the previous studies targeted auditory (speech) processing. As the core part of language processing is modality independent, we hypothesize that entrainment also occurs in visual modality. Moreover, theoretical linguistics proposed the single-engine hypothesis, that is, phrases/sentences and words have common hierarchical structures built by a fundamental syntactic operation, Merge (Halle & Marantz, 1993). A recent MEG study further supports this proposal, by showing that phrases and affixed words in English had similar cortical activations (Flick et al., 2018; see Gwilliams, 2020 for a theoretical review). However, it is difficult to conduct experiments on entrainment for words because the previous studies targeted a fusional language (e.g., English), of which morphemes tend to have multiple grammatical features and are less agglutinative, or an analytic language (e.g., Mandarin), which has almost no inflectional affixes. Therefore, experimental studies for typologically-different agglutinative languages (e.g., Japanese), which tend to have one grammatical feature for one morpheme, are critical to test whether phrases/sentences and words have common hierarchical structures built by Merge. In the present EEG experiment, we examine

whether words have hierarchical structures similar to phrases/sentences, by investigating an agglutinative language, Japanese. Furthermore, we used auditory and visual linguistic stimuli to test the modality independence of language processing. Here, using the isochronous rhythmic presentation (250 ms = 4 Hz) of morphemes (Figure 17), we examine whether the cortical entrainments are observed for words. We presented Japanese verbs with two, three, and four morphemes. Here we hypothesize that the rhythmic presentation of morphemes will elicit cortical entrainments that reflect the structure of words, and that both auditory and visual presentations of morphemes elicit similar patterns of cortical entrainments (Table 14).

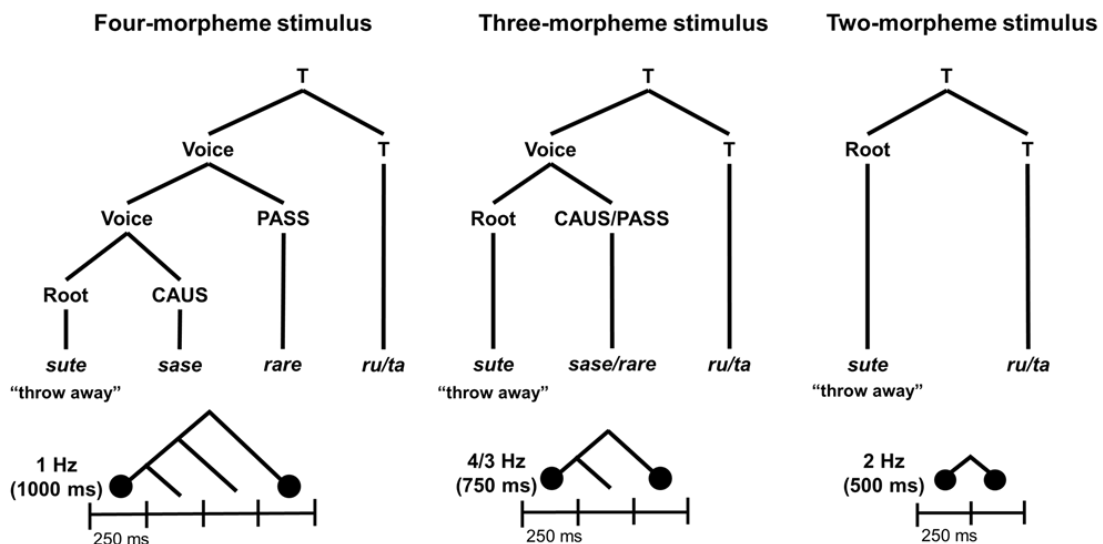


Figure 17. Schematic structures of word stimuli

Notice that these structures above are abbreviated that only depicts morphemes which have phonological/written form. The black filled circles denote the initial and final morphemes in a word. CAUS: causative, “sase”; PASS: passive, “rare”; T: tense, “ru”, non-past, “ta”, past.

Table 14. Prediction for the evoked power of the frequency of interest

Frequency of interest	
1 Hz	Four > Three, Two, and Random
4/3 Hz	Three > Four, Two, and Random
2 Hz	Two > Four, Three, and Random
4 Hz	Four = Three = Two = Random
8/3 Hz (resonance frequency)	Three > Four, Two, and Random
3 Hz (resonance frequency)	Four > Three, Two, and Random

5.1.1 Introduction

The two ERP experiments 1 & 2 reported in Chapters 3 & 4 solved the Japanese word processing theory's inconsistency. In addition, the results are consistent with the theory for English, Greek, and other languages. Moreover, the results demonstrated that the single mechanism composition theory is the plausible theory suggesting that the neural mechanism of complex word processing is language universal (3.5. & 4.5.1.; Introduction aim (1)). Besides, we demonstrated how words' complexity (structure) affects complex word processing, and the results suggested that the structure of words affects neural activation of the lexical processing, the N400 (4.5.2.). However, there were two experimental limitations to the ERP experiments. First, the structures of words used in experiment 2 are not consistent among

the stimuli by their structures resulting in the effect of the structure on the brain activation may have been confounded (e.g., left-branching [[*hoso*]-*naga*]]-[*sa*] and prefixed [*ko*]-[[*daka*]-[*sa*]]; Tagawa, 2008). Second, the (de-)adjectival stimuli are difficult to control their complexity due to the restriction in derivation.

In this experiment, we employ verbal stimuli whose structure is left-branching to solve these experimental problems with experiment 2. Specifically, by presenting the stimuli morpheme-by-morpheme in an isochronous rhythm, we demonstrate that the complexity of words is reflected in brain activation, the cortical entrainment. Moreover, comparing the previous studies, which demonstrate the synchronization between brain activation and structures of the sentences/phrases, we demonstrate a common neural mechanism in processing between words and sentences/phrases (Introduction aim (2): the single-engine hypothesis). Besides using the auditory and visual modality for stimuli, we demonstrate modality-independency of language processing.

5.2 Experiment

5.2.1 Participants

For the auditory experiment, we recruited 22 native Japanese speakers with normal or corrected-to-normal vision were recruited from Kyushu University in Japan (12 females, mean age $\pm$ *S.D.*: 21.0 ± 3.3 yrs.). All participants were classified as right-handed based on the Edinburgh handedness inventory (Oldfield, 1971: mean $\pm$ *S.D.*; 91.3 ± 11.6). Two participants were excluded from further analysis because of excessive artifacts in the EEG data. For the visual experiment, we recruited 27 native Japanese speakers with a normal or corrected-to-normal vision (12 females, mean age $\pm$ *S.D.*: 20.7 ± 3.3 yrs.). All participants were classified as right-handed based on the Edinburgh handedness inventory (Oldfield, 1971: mean $\pm$ *S.D.*; 91.3 ± 11.3). One participant was excluded from further analysis because of excessive artifacts in the EEG data. Fifteen participants joined both auditory and visual experiments. All participants provided written informed consent before participating in the study. Approval for the experiments was obtained from the institutional review board of the Faculty of Humanities, Kyushu University.

5.2.2 Stimuli

Using the Balanced Corpus of Contemporary Written Japanese (BCCWJ) (Maekawa et al., 2013), we selected 48 highly frequent transitive verbs with two syllables (moras) in their verbal roots (e.g., *sute-ru*, throw away). We created the following three experimental conditions in both auditory and visual experiments: two-morpheme, three-morpheme, and four-morpheme conditions. Stimuli under the **two-morpheme** condition contained a verbal root and a non-past or past tense marker, resulted in two stimuli for each root (e.g., “*sute-ru*” or *sute-ta*”; “捨て-た”; throw away or threw away). In the **three-morpheme** condition, a causative voice “-*sase-*” or a passive voice “-*rare-*” was inserted between the verbal root and a non-past/past tense marker, resulted in four stimuli for each root (e.g., “*sute-sase-ru*”, “*sute-rare-ru*”, “*sute-sase-ta*,” or “*sute-rare-ta*”; “捨て-させ-る”, “捨て-られ-る”, “捨て-させ-た”, or “捨て-られ-た”; make something throw away, get something throw away, made something throw away, or got something throw away, respectively). In the **four-morpheme** condition, both causative and passive voices were inserted before a tense marker, resulted in

two stimuli for each root (e.g., “*sute-sase-rare-ru*” and “*sute-sase-rare-ta*”; “捨て-させ-られ-る”, “捨て-させ-られ-た”; make and get something throw away and made and got something throw away respectively). Note that the causative-passive order is natural, while the passive-causative order is unacceptable in Japanese (Shibatani, 1990). We made a total of 384 stimuli (48 roots × 8 forms). Each stimulus had a left-branching structure and consisted of at least two morphemes (i.e., root, voice, or tense) (Figure 17). Notice that the structures in Figure 17 are abbreviated to show the phonological or written form of morphemes. Besides, the “Root” indicates an inflectional root as a morpheme but not the “√ROOT” commonly used in the distributed morphology.

We further created nonce word stimuli by changing one syllable of each verbal root, voice, or tense morphemes. For the verbal root of nonce word stimuli, the second syllable was replaced by “*mu*” or “*ru*” (e.g., **sumu-ru/ta*,” **捨む-る/た*” or **suru-ru/ta*,” *捨る-る/た*”). For the voices, “-*sase-*” was changed to “*mase*” or “*same*”, and “-*rare-*” was changed to “*rame*” or “*mare*” (e.g., **sute-mase-ru/ta*,” **sute-same-ru/ta*,” **sute-rame-ru/ta*,” and **sute-mare-ru/ta*”; **捨て-ませ-る/た*,” **捨て-さめ-*

る/た,” **“捨て-らめ-る/た,”* or **“捨て-まれ-る/た,”* respectively). For the tense markers, “-ru” or “-ta” was changed to “bu” or “gu” (e.g., **“sute-gu”* or **“sute-bu”*; **“捨て-ぐ”* or **“捨て-ぶ”*).

For the auditory experiment, all stimuli including nonce words were synthesized using the VOICEVOX version 0.11.4 (<https://voicevox.hiroshiba.jp/>, the male voice, Kurono Takehiro). Each morpheme was recorded separately to exclude a prosodic contour over the words. The durations of synthesized syllables were 150–250 ms (mean duration 230 ms) and were adjusted to 250 ms by padding silence at the end. For the visual experiment, the displayed fonts were Meiryo and Arial for Japanese characters and other letters, respectively. The stimulus presentation of the two experiments were controlled by the Presentation software (version 23.0: <https://www.neurobs.com/>).

5.2.3 Procedure

The experiment was organized into seven blocks (**No. of morpheme**): Six experimental blocks and one random (control) block. Each block consisted of 30 trials, and each trial consisted of sequential presentation of

24 morphemes. We presented two blocks for each of the three conditions: The **four-morpheme** condition (six verbs per trial), **three-morpheme** condition (eight verbs per trial), and **two-morpheme** condition (12 verbs per trial). For the control block, we used the **random** condition consisted of eight verbs: two stimuli from the Four-morpheme condition, four stimuli from the three-morpheme condition, and two stimuli from the two-morpheme condition. The word stimuli were presented morpheme-by-morpheme at 250 ms intervals. The order of stimuli in the random condition was pseudo-randomized within each trial. Thus, we presented 210 trials to each participant in each of the auditory and visual experiments. The order of the blocks was counterbalanced based on a Latin Square design by a participant.

Based on the previous neurolinguistic studies (Ding et al., 2017; Lo et al., 2022), we predict that each condition has a particular peak of frequencies corresponding to the number of morphemes (frequency of interest): A 1 Hz (1000 ms) peak for the four-morpheme condition, a $4/3$ Hz (750 ms) peak for the three-morpheme condition, and a 2 Hz (500 ms) peak for the two-morpheme condition. Also, the resonance frequencies for each experimental condition (e.g., $8/3$ Hz in the three-morpheme condition) will be observed.

We further predict 4 Hz (250 ms) peak, which reflects the inter-stimulus interval among all conditions. Participants conducted a deviation detection task. The participants were asked to answer whether a nonce word (deviation) was presented in each trial by pressing one of the two buttons (Cedrus, RB740). In each block, 24 trials without a nonce word and six trials with a nonce word were intermixed and presented in random order. The response buttons were counterbalanced among the participants. In the auditory experiment, morphemes with two syllables (i.e., root, causative voice, and passive voice) or one syllable (i.e., tense) were presented sequentially, while in the visual experiment, morphemes with two letters (i.e., root, causative voice, and passive voice) or one letter (i.e., tense) were presented sequentially with a duration of 250 ms.

During the experiment, the participants were seated in an electromagnetic shield room with a monitor positioned approximately 130 cm in front of them. The participants were instructed to judge whether or not the displayed letter strings were words. They completed more than two practice trials, including feedback before the experiment. In addition, each participant took a short break every five minutes during the experimental

blocks.

5.2.4 EEG data acquisition

We acquired the EEG by fitting the participants with an elastic cap with 64 embedded Ag/AgCl electrodes (actiCAP, Brain Products; Neurofax EEG-1200, Nihon Kohden). Two additional Ag/AgCl electrodes were placed below and to the left of the left eye to monitor vertical and horizontal eye movements. The linked earlobes served as a reference. All electrodes' impedances were maintained below 10 k Ω throughout the experiment. The ERPs were amplified with a bandpass of 0.05–58 Hz and sampled at 1,000 Hz.

5.2.5 Data analysis

We excluded the following trials from behavioral and EEG analysis: Trials with a nonce word, trials with incorrect responses, and trials with reaction times exceeding ± 2.5 *S.D.* from the average (Baayen, 2008). We considered a p-value of less than 0.05 statistically significant in the analyses. We applied the Greenhouse-Geisser correction for all effects involving more

than one degree of freedom (Greenhouse & Geisser, 1959, Picton et al., 2000) and corrected the multiple comparisons using Shaffer's modified sequentially rejective Bonferroni procedure (Shaffer, 1986). For significant main effects with more than three levels, we conducted post-hoc pairwise comparisons using paired *t*-tests. We also ran simple effects tests for significant interactions. We reported the corrected degrees of freedom, corrected *p*-values, and effect sizes (generalized eta squared [η^2_G] for ANOVA and Cohen's *d* for *t*-tests; Cohen, 1992). In addition, Type III sums of squares were applied when the analyzed data was unbalanced.

5.2.5.1 Behavioral data analysis

We conducted a two-way repeated-measures analysis of variance (rANOVA: **Modality** [**Auditory** and **Visual**] and **No. of Morpheme** [**Four-morpheme**, **Three-morpheme**, **Two-morpheme**, and **Random**]) analyses for the accuracy and reaction times.

5.2.5.2 EEG data analysis

We used MNE-Python for the EEG data analysis (Gramfort et al., 2013)

and applied independent component analysis (ICA) to de-noise the eyeblinks and ECGs. We computed EEG averages for a 6000 ms time window under all conditions. The baseline was set to 100 ms before the stimulus onset.

5.2.6 Data processing

We excluded trials with incorrect responses and large artifacts exceeding 500 μV from further analysis (approximately 7.51% of the data).

To avoid the transient EEG response to the stimulus onset, we removed the initial three seconds from each trial and analyzed the last 3 seconds (3–6 s: Ding et al., 2016). We assessed synchrony from 0.5–5 Hz with a frequency resolution of $1/3 = 0.3333$ Hz. The EEG data in analyzed trials were converted into the frequency domain by using the discrete Fourier transformation (DFT).

Evoked power reflected the power of EEG responses that were synchronized to the stimulus input. In equation (9) shown below, the DFT of the response in trial n was denoted as $X_n(f)$, and $X_n(f)$ was a complex-value Fourier coefficient. Thus, evoked power was the summation of complex-value Fourier coefficient of trials averaged over the total number of trials N .

$$E(f) = \frac{|\sum_n X_n(f)|^2}{N} \quad (9)$$

The statistical significance of the cortical entrainment at a target frequency was tested for evoked power. To remove the 1/f trend of evoked power, the spectrum was normalized by the neighboring four frequency bins (two bins on each side, which is equivalent to $\pm 2/3$ Hz) via Equation (10) adapted from Ding et al., (2017), where w denotes frequencies around the target frequency f .

$$E_n = \frac{E(f)}{\sum_w E(w)}, |w - f| < \frac{2}{3} \text{ Hz}, w \neq f \quad (10)$$

For statistical analyses, we conducted a one-way rANOVA for each measure at each frequency of interest: 1 Hz, 3/4 Hz, 2 Hz, and 4 Hz. In addition, to estimate normalized evoked powers of the resonance frequencies of the three-morpheme and four-morpheme conditions (i.e., 8/3 Hz and 3 Hz) were analyzed in the same manner.

5.3 Results

5.3.1 Behavioral results

The mean accuracy of the deviation detection task was 94.1% (**Four-morpheme**: $94.2 \pm 3.14\%$, **Three-morpheme**: $94.5 \pm 4.74\%$, **Two-**

morpheme: $94.3 \pm 4.75\%$, **Random:** $93.7 \pm 8.86\%$; **Auditory:** $94.5 \pm 9.41\%$, **Visual:** $93.7 \pm 3.14\%$), indicating that the participants correctly performed the task.

For the accuracy of a two-way rANOVA (**Modality** [**Auditory** and **Visual**] $\times$ **No. of Morpheme** [**Four-morpheme**, **Three-morpheme**, **Two-morpheme**, and **Random**]), there were no significant main effects and interaction ($ps > .13$) [**Modality:** $F(1, 44) = 0.528, p = .47, \eta^2_G = .0049$; **No. of Morpheme:** $F(1, 2.74) = 0.0491, p = .98, \eta^2_G = .0007$; **interaction:** $F(1, 120.57) = 1.88, p = .14, \eta^2_G = .025$].

The reaction times showed the significant main effects for modality and no. of morpheme [**Modality:** $F(1, 44) = 62.6, p < .0001, \eta^2_G = .55$, **No. of Morpheme:** $F(30.86, 2.17) = 11.2, p < .0001, \eta^2_G = .039$]. The visual modality showed longer reaction times than the auditory. Multiple comparisons for No. of Morpheme showed that the random condition yielded shorter reaction times than the other conditions, suggesting that the random condition required lower processing loads than the other experimental conditions [**vs. Four-morpheme:** $t(44) = 5.28$, corrected $p < .0001$, Cohen's $d = 0.27$; **vs. Three-morpheme:** $t(44) = 5.49$, corrected $p < .0001$, Cohen's

$d = 0.28$; **vs. Two-morpheme**: $t(44) = 4.70$ corrected $.p = .0001$, Cohen's $d = 0.37$]. We found no significant difference among the other conditions ($ps > .40$) [**Four-morpheme vs. Three-morpheme**: $t(44) = 0.0234$, corrected $.p = .98$, Cohen's $d < 0.001$; **Four-morpheme vs. Two-morpheme**: $t(44) = 1.14$, corrected $.p = .52$, Cohen's $d = 0.09$; **Three-morpheme vs. Two-morpheme**: $t(44) = 1.52$, corrected $.p = .41$, Cohen's $d = 0.10$]. The interaction between modality and no. of morpheme was not significant [$F(2.17, 95.61) = 1.03$, $p = .37$, $\eta^2_G = .0037$].

5.3.2 EEG results

5.3.3 Auditory results

In the experimental conditions, we found clear entrainment peaks for the normalized evoked power: **1 Hz** for the four-morpheme condition, **4/3 Hz** for the three-morpheme condition, and **2 Hz** for the two-morpheme condition, respectively (Figure 18). Furthermore, the evoked powers for the cortical entrainments were significantly stronger than the neighboring frequencies (Figure 19), replicating the previous studies (Ding et al., 2016; Ding et al., 2017; Lo et al., 2022). In contrast, the random condition elicited

only one entrainment peak at 4 Hz, indicating that the random presentation of morphemes only induced cortical entrainments corresponding to exogenous events (i.e., 4 Hz auditory stimuli). Besides, as resonant frequencies, 2 Hz and 3 Hz peaks were also observed for the four-morpheme condition, and an 8/3 Hz peak was observed for the three-morpheme condition.

We observed the main effect of the condition at 1 Hz [$F(1.95, 37.01) = 19.9, p < .0001, \eta^2_G = .45$]. Post-hoc comparison showed that the four-morpheme condition had a significantly stronger peak than the other conditions [**vs. Three-morpheme**: $t(19) = 5.92$, corrected $p < .0001$, Cohen's $d = 2.06$; **vs. Two-morpheme**: $t(19) = 3.52$, corrected $p = .0068$, Cohen's $d = 0.55$; **vs. Random**: $t(19) = 6.04$, corrected $p < .0001$, Cohen's $d = 1.37$] (Figure 19). We also observed the main effect of the condition at 4/3 Hz [$F(1.42, 26.94) = 25.5, p < .0001, \eta^2_G = .53$]. Post-hoc comparison showed that the three-morpheme condition had a significantly stronger peak than the other conditions [**vs. Four-morpheme**: $t(19) = 4.95$, corrected $p = .0003$, Cohen's $d = 1.82$; **vs. Two-morpheme**: $t(19) = 6.36$, corrected $p < .0001$, Cohen's $d = 2.02$; **vs. Random**: $t(19) = 5.29$, corrected $p = .0001$, Cohen's d

= 1.73]. We found the main effect of the condition at 2 Hz as well [$F(1.69, 32.08) = 24.3, p < .0001, \eta^2_G = .42$]. Post-hoc comparison showed that the two-morpheme condition had a significantly stronger peak than the other conditions [**vs. Four-morpheme**: $t(19) = 4.18$, corrected $p = .0015$, Cohen's $d = 0.99$; **vs. Three-morpheme**: $t(19) = 5.05$, corrected $p = .0002$, Cohen's $d = 1.59$; **vs. Random**: $t(19) = 6.52$, corrected $p < .0001$, Cohen's $d = 1.68$].

On the contrary to the other frequencies of interest, the 4 Hz peak did not show significant difference among all conditions [$F(2.22, 42.24) = 2.88, p = .062, \eta^2_G = .055$; **Four-morpheme vs. Three-morpheme**: $t(19) = 1.22$, corrected $p = .47$, Cohen's $d = 0.31$; **Four-morpheme vs. Two-morpheme**: $t(19) = 1.48$, corrected $p = .46$, Cohen's $d = 0.37$; **Three-morpheme vs. Two-morpheme**: $t(19) = 0.182$, corrected $p = .86$, Cohen's $d = 0.04$; **Four-morpheme vs. Random**: $t(19) = 2.89$, corrected $p = .056$, Cohen's $d = 0.60$; **Three-morpheme vs. Random**: $t(19) = 1.52$, corrected $p = .44$, Cohen's $d = 0.33$; **Two-morpheme vs. Random**: $t(19) = 1.94$, corrected $p = .20$, Cohen's $d = 0.33$].

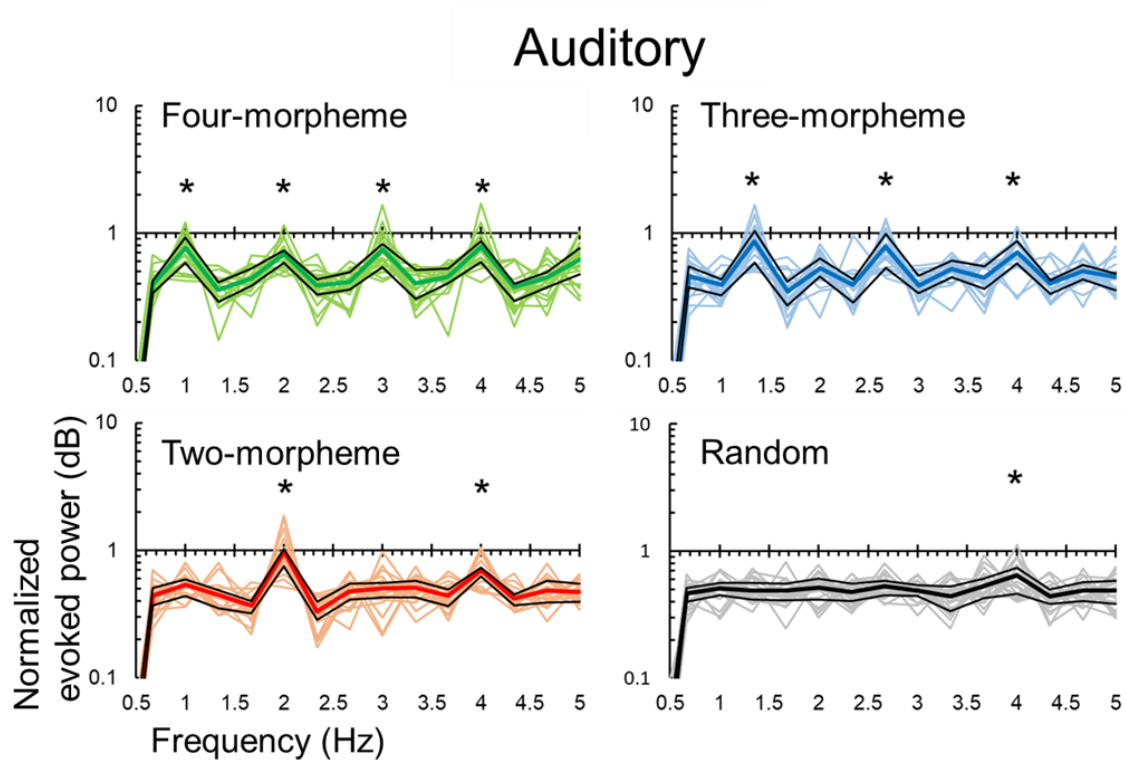


Figure 18. Normalized evoked power (grand average over all channels) of auditory stimuli for four conditions (log scale) Four-morpheme (green), three-morpheme (blue), two-morpheme (red), and random (black) conditions. Bold line shows the grand average over subjects, while the two thin black lines show the 75th and 25th percentiles. Thin lines show the data from individual subjects. Asterisks denote frequency bins showing stronger power than the mean power of a neighboring 2/3 Hz on each side (corrected $p < 0.05$, paired two-sided t -test).

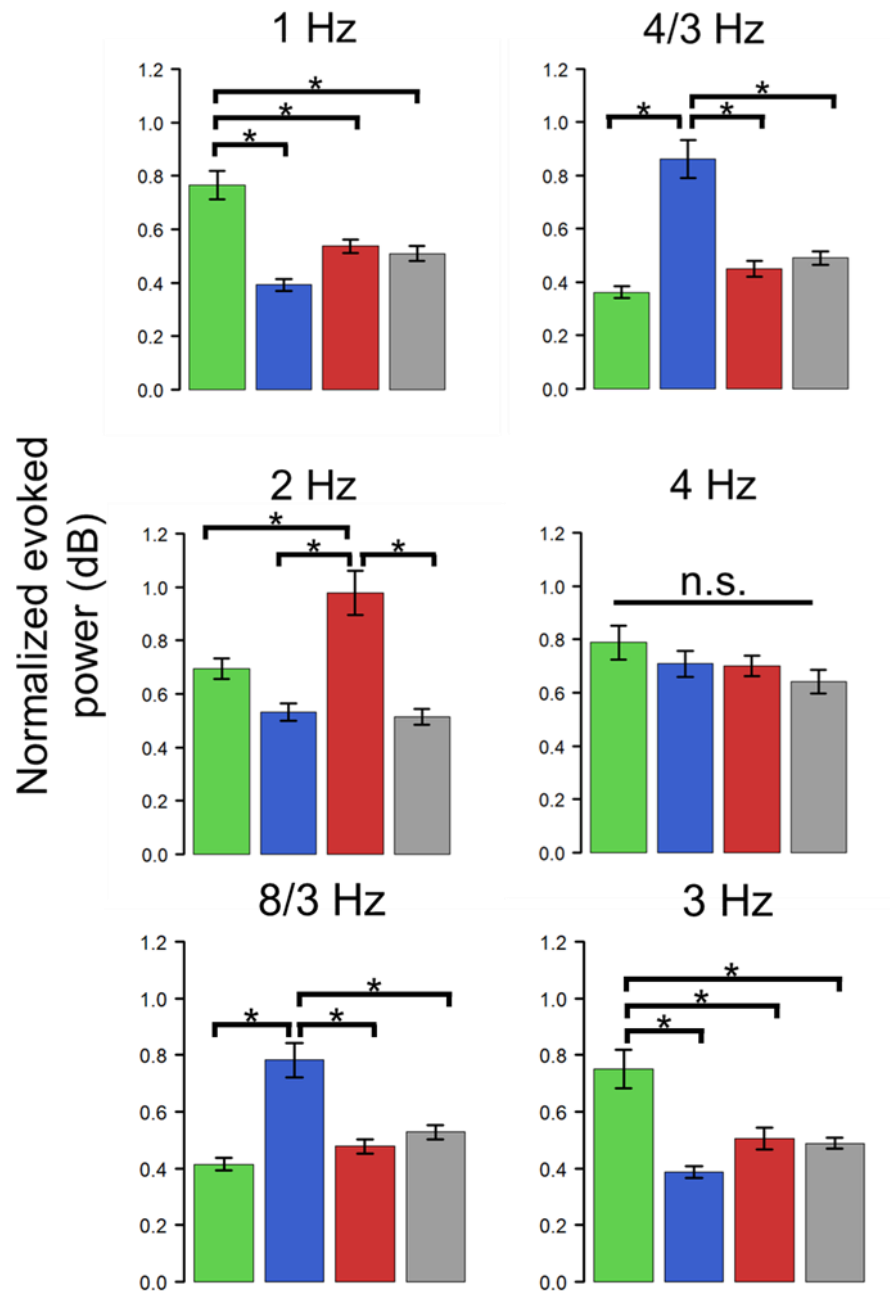


Figure 19. The spectral activity across four conditions for normalized evoked power in the auditory experiment
Frequency of interest for experimental conditions: Four-morpheme (green), 1 Hz; three-morpheme (blue), 4/3 Hz; two-morpheme (red), 2 Hz. Error bars denote standard error of the mean. Significance code: ‘*’ 0.05.

For the resonance frequency, we observed the main effect of the condition at 8/3 Hz [$F(1.65, 31.28) = 19.0, p < .0001, \eta^2_G = .4286$]. Post-hoc comparison showed a statistically significant strong peak in the comparison of the three-morpheme condition and each of the others [**vs. Four-morpheme**: $t(19) = 4.90$, corrected $p = .0003$, Cohen's $d = 1.79$; **vs. Two-morpheme**: $t(19) = 5.23$, corrected $p = .0003$, Cohen's $d = 1.47$; **vs. Random**: $t(19) = 4.18$, corrected $p = .0015$, Cohen's $d = 1.23$]. The main effect of the condition at 3 Hz was also significant [$F(1.51, 28.65) = 12.0, p = .0005, \eta^2_G = .35$]. Post-hoc comparison showed a statistically significant strong peak in the comparison of the four-morpheme condition and the three-morpheme and random conditions [**vs. Three-morpheme**: $t(19) = 4.90$, corrected $p = .0006$, Cohen's $d = 1.61$; **vs. Random**: $t(19) = 3.94$, corrected $p = .0027$, Cohen's $d = 1.17$]. Comparison between the four-morpheme and the two-morpheme condition was marginally significant [**vs. Two-morpheme**: $t(19) = 2.61$, corrected $p = .052$, Cohen's $d = 0.99$].

Statistical comparisons at resonance frequencies of interest (i.e., 8/3 Hz and 3 Hz) in the auditory experiment are illustrated in Figure 19. For normalized evoked power, the resonance frequencies of interest for each experimental

condition were marginally stronger than the other three conditions (corrected $ps < .053$).

5.3.4 Visual results

Same as the results of the auditory experiment, the frequencies of interest showed salient peaks for each experimental condition (Figure 20). Statistical comparisons at each frequency of interest are illustrated in Figure 21. These results demonstrated that the entrainment induced by hierarchical structures was modality-independent. Besides, the peaks of resonance frequencies were observed: 2 Hz and 3 Hz peaks for the four-morpheme condition, as well as 8/3 Hz peak for the three-morpheme condition.

For normalized evoked power, the main effect of the condition at 1 Hz was significant [$F(1.58, 39.44) = 38.3, p < .0001, \eta^2_G = .52$]. Post-hoc comparison showed the four-morpheme condition had a significantly stronger peak than the other conditions [vs. Three-morpheme: $t(25) = 7.25$, corrected $p < .0001$, Cohen's $d = 1.80$; vs. Two-morpheme: $t(25) = 6.55$, corrected $p < .0001$, Cohen's $d = 1.96$; vs. Random: $t(25) = 6.37$, corrected $p < .0001$, Cohen's $d = 2.02$]. The main effect of the condition was also found

for the 4/3 Hz peak [$F(1.28, 31.9) = 43.2, p < .0001, \eta^2_G = .56$]. Post-hoc comparison showed the three-morpheme condition had a significantly stronger peak than the other conditions [vs. Four-morpheme: $t(25) = 6.97$, corrected $p < .0001$, Cohen's $d = 2.04$; vs. Two-morpheme: $t(25) = 7.29$, corrected $p < .0001$, Cohen's $d = 1.91$; vs. Random: $t(25) = 6.23$, corrected $p < .0001$, Cohen's $d = 1.66$]. We found the main effect of the condition at 2 Hz as well [$F(2.03, 50.67) = 40.5, p < .0001, \eta^2_G = .52$]. Post-hoc comparison showed the two-morpheme condition had a significantly stronger peak than the other conditions [vs. Four-morpheme: $t(25) = 3.96$, corrected $p = .0011$, Cohen's $d = 0.87$; vs. Three-morpheme: $t(25) = 7.07$, corrected $p < .0001$, Cohen's $d = 1.96$; vs. Random: $t(25) = 7.20$, corrected $p < .0001$, Cohen's $d = 2.17$].

In contrast to the other frequencies of interest, the 4 Hz peak did not show significant difference among all conditions [$F(2.49, 62.15) = 2.52, p = .0771, \eta^2_G = .026$; Four-morpheme vs. Three-morpheme: $t(19) = 1.35$, corrected $p = .57$, Cohen's $d = 0.23$; Four-morpheme vs. Two-morpheme: $t(19) = 2.55$, corrected $p = .10$, Cohen's $d = 0.44$; Three-morpheme vs. Two-morpheme: $t(19) = 2.32$ corrected $p = .10$, Cohen's $d = 0.23$; Four-morpheme

vs. Random: $t(19) = 0.872$, corrected $p = .783$, Cohen's $d = 0.18$; Three-morpheme vs. Random: $t(19) = 0.240$, corrected $p = .81$, Cohen's $d = 0.04$; Two-morpheme vs. Random: $t(19) = 1.54$, corrected $p = .41$, Cohen's $d = 0.25$].

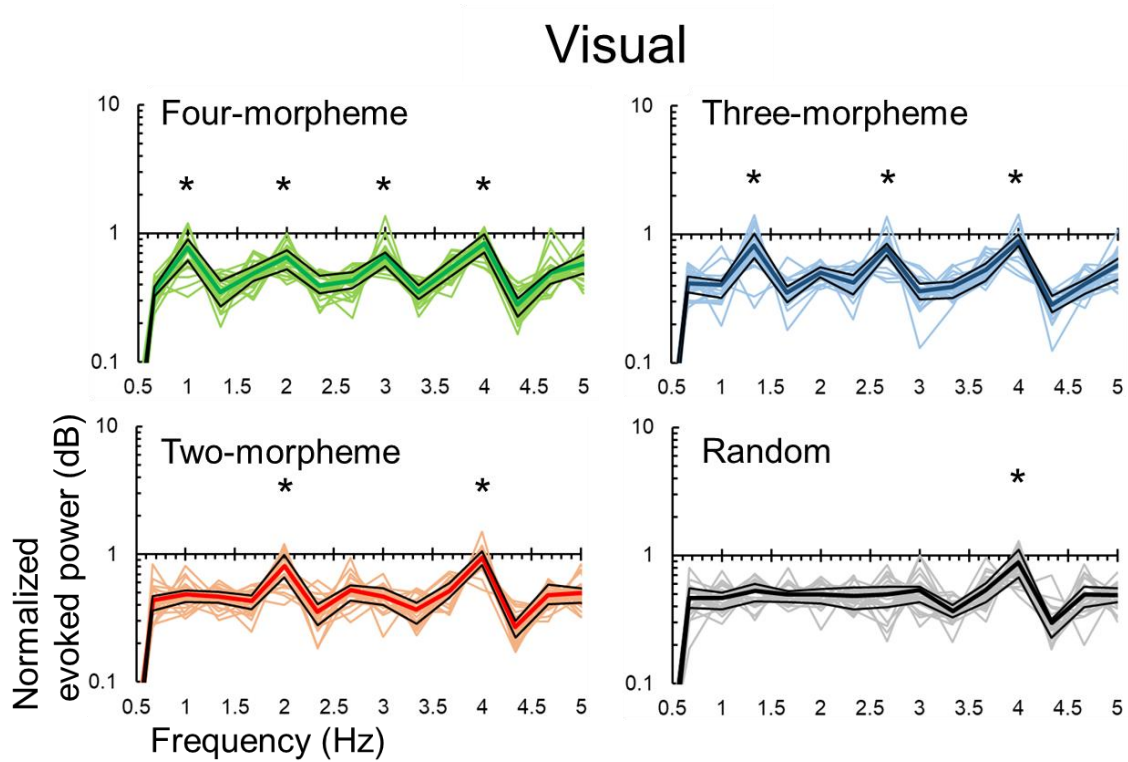


Figure 20. Normalized evoked power (grand average over all channels) of visual stimuli for four conditions (log scale)

Four-morpheme (green), Three-morpheme (blue), Two-morpheme (red), and Random (black). Bold line shows the grand average over subjects, the two thin black lines delimit the 75th and 25th percentiles. Thin lines show the data from individual subjects. Star denotes frequency bins showing power stronger than the mean power of a neighboring 2/3 Hz on each side (corrected $p < 0.05$, paired two-sided t -test).

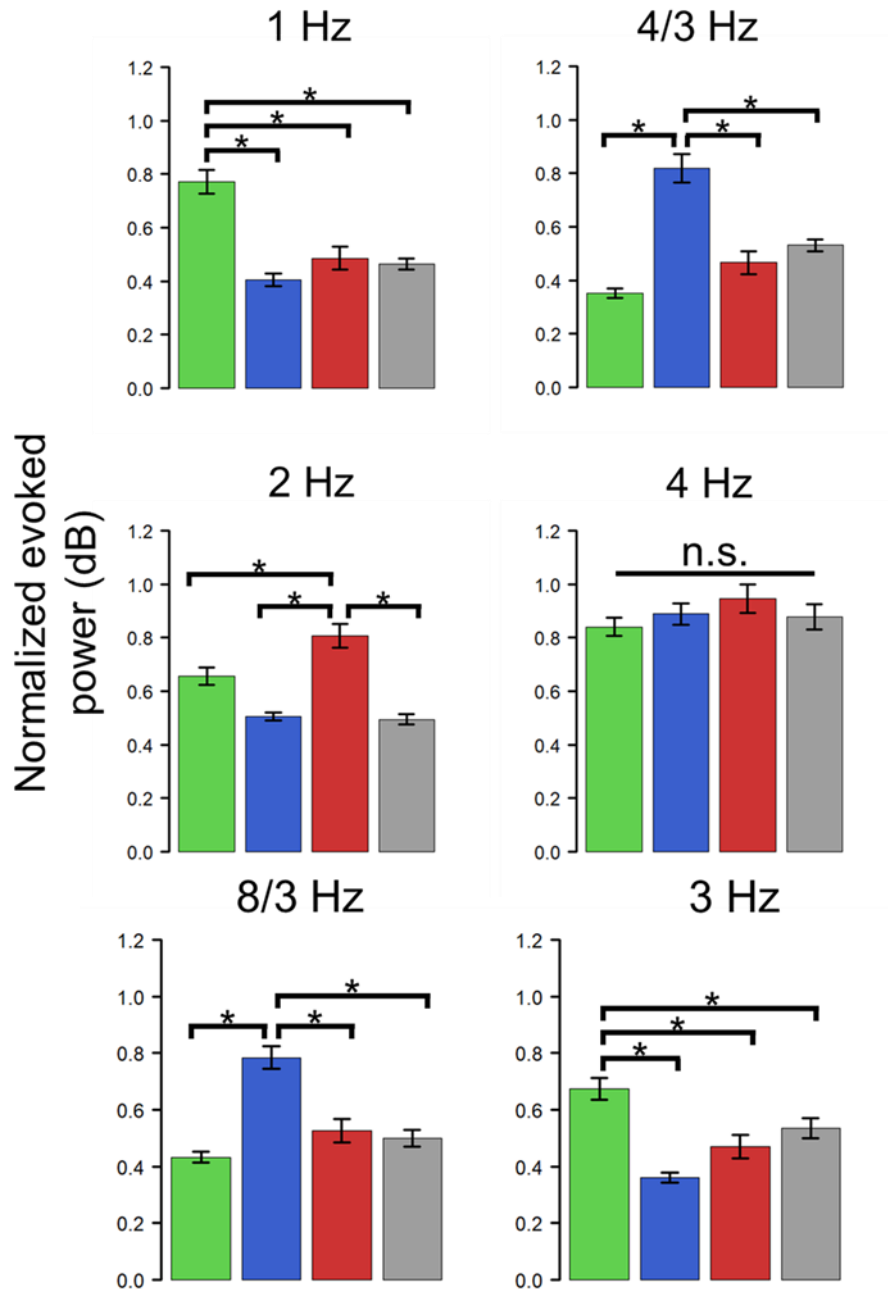


Figure 21. The spectral activity across four conditions for normalized evoked power in the visual experiment
Frequency of interest for experimental conditions: Four-morpheme (green), 1 Hz; three-morpheme (blue), 4/3 Hz; two-morpheme (red), 2 Hz. Error bars denote standard error of the mean. Significance code: '*' 0.05.

For the resonance frequency, we observed the main effect of the condition at 8/3 Hz [$F(1.69, 42.33) = 48.0, p < .0001, \eta^2_G = .59$. Post-hoc comparison showed a statistically significant strong peak in the comparison of the three-morpheme condition and each of the others [**vs. Four-morpheme:** $t(25) = 8.53$, corrected $p < .0001$, Cohen's $d = 2.33$; **vs. Two-morpheme:** $t(25) = 7.40$, corrected $p < .0001$, Cohen's $d = 2.01$; **vs. Random:** $t(25) = 7.01$, corrected $p < .0001$, Cohen's $d = 2.33$]. In addition, the main effect of the condition at 3 Hz was significant [$F(1.94, 48.55) = 32.2, p < .0001, \eta^2_G = .47$]. Post-hoc comparison showed a statistically significant strong peak in the comparison of the four-morpheme condition and each of the others [**vs. Three-morpheme:** $t(25) = 7.22$, corrected $p < .0001$, Cohen's $d = 2.19$; **vs. Two-morpheme:** $t(25) = 6.42$, corrected $p < .0001$, Cohen's $d = 1.48$; **vs. Random:** $t(25) = 4.65$, corrected $p = .0003$, Cohen's $d = 1.23$].

Statistical comparisons at resonance frequencies of interest (i.e., 8/3 Hz and 3 Hz) in visual experiment are illustrated in Figure 21. For normalized evoked power, the resonance frequencies of interest for each experimental condition were significantly stronger than the other three conditions

(corrected $ps < .0004$).

5.3.5 Comparison between auditory and visual modalities

To compare the cortical entrainment between modalities, a two-way rANOVA (**Modality [Auditory and Visual] $\times$ No. of Morpheme [Four-morpheme, Three-morpheme, Two-morpheme, and Random]**) was conducted for the peaks of each frequency of interest (i.e., 1, 4/3, 2, 4 Hz).

The main effect of modality was not significant in the three frequencies of interest [**1 Hz**: $F(1, 45) = 0.0009, p = .98, \eta^2_G < .0001$; **4/3 Hz**: $F(1, 45) = 0.652, p = .42, \eta^2_G = .0029$; **2 Hz**: $F(1, 45) = 1.93, p = .17, \eta^2_G = .016$]. In contrast, the peak for 4 Hz was significantly stronger in the visual modality [$F(1, 45) = 17.2, p < .01, \eta^2_G = .18$], suggesting that the visual processing elicits more cortical activation.

The main effect of no. of morpheme was significant in the 1 Hz, 4/3 Hz, and 2Hz [**1 Hz**: $F(0.9911, 1.58) = 64.1, p < .01, \eta^2_G = .51$; **4/3 Hz**: $F(1.716, 1.36) = 67.4, p < .01, \eta^2_G = .55$; **2 Hz**: $F(1.871, 1.92) = 75.7, p < .01, \eta^2_G = .51$], whereas the main effect for 4 Hz was not [$F(0.0157, 2.91) = 1.09, p = .35, \eta^2_G = .0098$]. These results suggested that the peaks of frequencies of

interest (1 Hz, 4/3 Hz, and 2 Hz) were specific for each condition, while the 4 Hz peak, which reflected the isochronous stimulus presentation (i.e., 250 ms = 4 Hz), was not. The interaction was also not significant [**1 Hz**: $F(1.58, 71.19) = 0.510, p = .56, \eta^2_G = .0083$; **4/3 Hz**: $F(1.36, 61.15) = 0.222, p = .71, \eta^2_G = .0039$; **2 Hz**: $F(1.92, 86.61) = 1.33, p = .27, \eta^2_G = .018$; **4 Hz**: $F(2.91, 130.75) = 1.78, p = .16, \eta^2_G = .016$].

To compare the cortical entrainment between modalities, an additional one-way ANOVA for modality was conducted for peak of each frequency of interest (resonance frequencies: 8/3 Hz, and 3 Hz). For the peaks of 8/3 Hz, the main effect of modality was not significant [$F(1, 44) = 0.0005, p = .98, \eta^2_G < .0001$]. For the peaks of the 3 Hz, the main effect of modality was marginally significant [$F(1, 44) = 1.03, p = .32, \eta^2_G = .023$].

5.4 Discussion

By investigating the normalized evoked power of EEG induced by polymorphemic words under the isochronous stimulus presentation (Figure 17), we obtained four striking results. First, we demonstrated that the cortical entrainment was induced by the hierarchical structure of words in the

auditory modality (Figure 18 & Figure 19), replicating the result of cortical entrainments related to the hierarchical structures of phrases and sentences (Ding et al., 2016; Ding et al., 2017; Lo et al., 2022). These results strongly supported the single-engine hypothesis, i.e., the hierarchical structures of words and those of sentences/phrases were processed by a common neural mechanism. Second, we demonstrated a similar result for both of the auditory and visual experiments, suggesting that the processing of words was modality-independent (Figure 20 & Figure 21). Third, the two-way rANOVA (**Modality** $\times$ **No. of Morpheme**) on reaction times showed that the auditory modality was significantly faster, indicating that the auditory modality had lower processing loads. The longer reaction times in the visual experiment may be caused by unnatural stimulus presentation, i.e., morpheme-by-morpheme visual presentation. Fourth, the two-way rANOVA (**Modality** $\times$ **No. of Morpheme**) on the peak amplitude of each frequency of interest (i.e., 1, 4/3, 2, 4 Hz) also showed that the random condition elicited a stronger peak for 4 Hz in the visual modality. These results further suggested that the auditory modality is a default language processing modality.

In the auditory experiments, we found the peaks of frequencies

corresponding to the hierarchical structure of the words in the experimental conditions. In contrast, the random condition only elicited the frequency of the isochronous stimulus presentation (Figure 18). These results support the single-engine hypothesis proposed in the distributed morphology. In the visual experiment, we further found similar results in the auditory experiment (Figure 20 & Figure 21). These results suggested that the neural mechanism of word processing is modality independent. These results were consistent with the results of Flick et al. (2018), which reported that suffixed words and phrases had a common neural mechanism.

One concern in the current experiments is the probabilistic predictability of succeeding morphemes. In this paradigm, verbal roots were followed by one of the two voices or two tense markers, therefore, the participants could easily predict upcoming linguistic unit(s) (i.e., morpheme). Due to the higher predictability of upcoming morphemes, the observed entrainments of words might reflect the prediction of morphemes but not the hierarchical structure of words. However, Ding et al. (2016) demonstrated that entrainments reflected the hierarchical structure but not predictability. In addition, Lo et al. (2022) also demonstrated that the cortical entrainments

reflected hierarchical structure, but not the sequential lexical information (word categories). Based on these studies, we think the entrainments in the experiment also reflected the hierarchical structures of words.

One of the most important issues in psycho/neurolinguistics is how the syntactic processing is mapped to the neural activations. As most of the previous studies dealt with the “mapping problem” targeted fusional or analytic languages (German: Meyer, 2018; English and Mandarin Chinese: Ding et al., 2016; Mandarin Chinese: Ding et al., 2017 and Lo et al., 2022), we must target typologically difference languages to elucidate the correspondence between a syntactic operation and a neural activity. For example, previous MEG/EEG studies examined the morphological processing of Japanese verbs and derived nouns and found similar morpho-syntactic processing to English (Ohta et al., 2019; Nakajima & Ohta, 2023). We should also examine more broader linguistic phenomena to extend further our understanding of the neural basis of language processing.

Whether or not words differ from sentences/phrases is another core problem in theoretical and experimental linguistics. For example, the distributed morphology proposed that words and sentences/phrases have a

common syntactic hierarchical structure (Halle & Marantz, 1993). In line with a recent MEG study that suggested the plausibility of this hypothesis (Flick et al., 2018), the current experiment further supports this hypothesis. The current experiments demonstrated that a word with a hierarchical structure was processed by the common neural mechanism of sentences and phrases. Moreover, we demonstrated that the processing mechanism of words is modality independent. Although we demonstrated modality independence of the cortical entrainments for the word stimuli, it is still unclear whether the larger linguistic units (i.e., phrases and sentences) also show the modality-independence.

Studies using more natural stimuli provided controversial results. For instance, Luo and Ding (2020) reported that the delta band peaks were not observed for the effect of the structure of sentences in non-metrical stories. As pointed out by Lo et al. (2022), whether the interpretation of the results of the present study can be applied to more natural (i.e., not time-locked) stimuli remained unclear. Furthermore, Meyer et al. (2020) proposed that the two forms of synchronized neural oscillations were required for speech processing: Entrainment induced by exogenous stimulus and intrinsic

synchronicity derived from functionalities in the generation of abstract representations and predictions. Therefore, we must conduct more studies using natural stimuli to fully understand the neural basis of language.

5.5 Conclusion

We tried to examine whether words were processed similarly as sentences or phrases. Moreover, we examined modality independency in language processing. We hypothesized a common neural mechanism for processing words and sentences/phrases. In addition, we also hypothesized a modality independency in language (word processing). In the EEG analysis, we demonstrated the cortical entrainment occurs in auditory word stimuli (Figure 18 & Figure 19), suggesting that existence of the common neural mechanism between words and sentences/phrases. The result of cortical entrainment was consistent with the studies of cortical entrainment of sentences/phrases in English and Mandarin Chinese (Ding et al., 2016; Ding et al., 2017; Lo et al., 2022), and experimentally supported the hypothesis of the distributed morphology. Besides, as well as experiment 1 & 2, this experiment broadened that the language universality in processing of

typologically different language. Moreover, we also demonstrated that modality independency by comparing the auditory and visual experiments (Figure 20 & Figure 21). In addition, the reaction times also suggested that the auditory stimuli elicited lower processing loads. Although further evidence is needed to elucidate a broader mechanism for the language processing, the results clearly demonstrated a common processing mechanism of word and sentences/phrases. By demonstrating that the cortical entrainments of Japanese verb were consistent with those of sentences/phrases, the experiment strongly supported the language universality of neural basis of language processing.

6. General Discussion

In this dissertation, we tried to elucidate the neural basis of morphologically complex words. In Chapter 1, we reviewed the three hypotheses for processing morphologically complex words and formulated the remaining gaps in this field.

Specifically, we investigated two experimental gaps repeated below.

- (1) To solve the inconsistent results of word processing in Japanese de-adjectival nouns (Chapters 3 & 4: the single mechanism composition theory).
- (2) To demonstrate a common neural mechanism between words and phrases/sentences. (Chapter 5: the single-engine hypothesis).

Chapter 2 introduced the neurophysiological methods and brain activations to interpret the results of this dissertation. The following sections discuss and summarize the experiment results.

6.1 Morphological decomposition of complex words

Neurolinguistic studies have reported that the single mechanism composition theory is plausible theory for the processing of the complex words (Fruchter et al., 2013). However, the studies of Japanese de-adjectival nouns (*-mi* type) offered inconsistent results (Hagiwara et al., 1999; Clahsen

& Ikemoto, 2012). Focusing on the morphological decomposition, N170, we aimed to demonstrate that the single mechanism composition theory is plausible for the de-adjectival (-*mi* type) nouns, and both nouns are treated as morphemes (Chapter 3: experiment 1). In addition, by using the *hiragana* script, we aimed to exclude unexpected effects on decomposition. In the decompositional stage, the result replicated the previous neurolinguistic studies. On the other hand, in the latter stage, lexical processing, the script showed an unexpected effect on it. These results are consistent with the previous report on mixed scripts in Japanese and do not affect decomposition (Ohta et al., 2019) but lexical processing (Nakano & Kishimoto, 2019). In the stage of decomposition, however, Lewis et al. (2011) reported that genuine and pseudo-complex words also evoke similar activation (e.g., “farm-er” and “broth-er”). Therefore, whether the two de-adjectival (-*mi* type) nouns are treated as genuine complex words was still being unclear.

6.2 Lexical processing of complex words

To fill the remaining gap in the lexical processing, we conducted further experiment to investigate whether the nouns are genuinely complex by

focusing on the latter stage of lexical processing, N400 (Chapter 4: experiment 2). Applying the masked priming paradigm, we aimed to demonstrate the two de-adjectival nouns (*-mi* type) are genuine complex words. We used the mixed script (i.e., *hiragana* and *kanji*) to avoid unexpected effects in the stage of lexical processing. In addition, to investigate the complexity of the complex words, we conduct additional analyses for other de-adjectival nouns (non-*mi* types).

In the mixed-effect (LME) model analysis, we demonstrated that the *-mi* type is a genuine complex word (4.3.1.). Moreover, we demonstrated that the LME showed two temporal clusters which correlate with the TP and N170, which suggests that the clusters correspond to the lexical access and recombination (Fruchter et al., 2013; Neophytou et al., 2018; Stockall et al., 2019). The result is consistent with the processing of the complex words in other languages, suggesting the neural basis complex word processing is language universal. In addition, we found dissimilarity between two nouns by the ERP and behavioral analyses, suggesting the difference in the experimental paradigm causes inconsistency among studies. This interpretation explains the inconsistent results between Hagiwara et al.

(1999) and Clahsen and Ikemoto (2012).

Besides, additional analyses for other de-adjectival nouns replicated the result of the *-mi* type (4.3.2.). In addition, the result demonstrated that the number of morphemes (complexity) is a crucial predictor of the lexical process. Moreover, the number of morphemes and TP showed clear distinction in the time window and their regression coefficients.

To summarize, the experiment1 & 2 elucidate and solve inconsistency of the derived nouns (Introduction aim (1)).

6.3 Cortical entrainment

In Chapter 5 (experiment 3), based on the results of experiments 1 & 2, we addressed how the processing of complex words is reflected (“mapped”) in neural activations (Introduction aim (2)). Employing a similar paradigm for the phrases/sentences processing from previous studies, we further addressed to demonstrate that the common neural mechanism of sentences and phrases processes the word with a hierarchical structure (the single-engine hypothesis; Halle & Marantz, 1993). In addition, we used the auditory and visual modalities as stimuli to demonstrate the modality-independency

in language processing.

The results of complex words replicated the previous studies of phrases and sentences, suggesting the common neural mechanism among linguistic units, words, phrases, and sentences. The result strongly suggested that validity of the single-engine hypothesis proposed in the distributed morphology. Moreover, the comparison between the auditory and visual experiments demonstrated that the processing of words was modality-independent. In addition, the behavioral and EEG results suggested that the default modality for language processing is the auditory modality. Therefore, the results solved the Introduction aim (2) by extension. Furthermore, if the single-engine hypothesis is correct that complex words and sentences are built by a common mechanism, experiment 3 provided strong experimental evidence for the hypothesis and a reasonable result from the perspective of simplicity. In addition, as well as the result of experiments 1 & 2, the results also demonstrated that the neural basis of processing complex words (or further phrases/sentences) is language universal.

7. Conclusions

In neurophysiological experiments on languages, connecting language theories and experiments has been one of the most general problems (Embick & Poeppel, 2015). In this dissertation, we tried to examine the linguistic theory and hypothesis through three experiments. We conducted neurophysiological experiments that took language theory into account and demonstrated the validity of theoretical linguistics through neurolinguistic experiments. Specifically, in this dissertation, we tried to fill two experimental gaps in processing morphologically complex words. The aim of this dissertation is as follows.

- (1) To solve the inconsistent results of word processing in Japanese de-adjectival nouns (Chapters 3 & 4 : the single mechanism composition theory).
- (2) To demonstrate a common neural mechanism between words and phrases/sentences. (Chapter 5: the single-engine hypothesis).

In Chapters 1 & 2, we provided a theoretical and experimental review of the processing of the morphologically complex word and formulated two experimental gaps as aims shown above.

In Chapter 3 (experiment 1), we conducted an ERP study to solve the inconsistent theories of word processing by investigating the neural profile

of visual complex word processing focusing on the former time stage, morphological decomposition (aim (1)). By using the transition probability (TP) as a probe, the experiment supported the plausibility of the single mechanism composition theory. Meanwhile, the result suggested that the unique mixed script in Japanese affects the latter time stage of processing uninterpretable. The result provided partial evidence of the aim (1).

In Chapter 4 (experiment 2), we conducted an additional ERP study to investigate the neural profile of the latter time stage, lexical processing. The result demonstrated that the Japanese de-adjectival nouns are similarly treated as morphemes, revealing that the single mechanism composition theory is plausible (3.4.). In addition, we also found the dissimilarity of two nouns, suggesting the inconsistency of the studies on two nouns is caused by the time stage of processing and differences in the experimental paradigm.

Summarizing the results of Chapters 3 & 4 (experiments 1 & 2), we solved the aim (1) by demonstrating that the two de-adjectival nouns are processed by the single mechanism composition theory (4.4.1.). Moreover, we demonstrated that the neural basis of morphologically complex word processing is language universal (Table 5). Besides, additional analyses for

the complexity (structure) of words suggested that, as well as the TP, the complexity is a crucial predictor for the lexical processing (4.4.2.). Furthermore, the complexity and TP modulate the activation of lexical processing differently in their time stages and regression coefficient.

Chapter 5 (experiment 3) addressed the aim (2), the single-engine hypothesis. Using the isochronous rhythmic presentation of hierarchical complex words morpheme-by-morpheme, we revealed that word stimuli elicit a similar activation of the phrases/sentences, as reported in previous neurolinguistic studies (Ding et al., 2016, 2017; Lo et al., 2022). Moreover, we showed that the activation is modality-independent. The result experimentally demonstrates that the neural basis of words and phrases/sentences has a common neural mechanism, the single-engine hypothesis proposed by Halle and Marantz (1993). In addition, as well as experiments 1 and 2, experiment 3 provided further evidence of universality in language processing.

There are four possible future directions. First, now that the hierarchical structure of words and phrases/sentences has a common and modality-independent neural mechanism in Japanese, we must investigate whether it

is replicative in the other languages. As experiment 1 suggested that the unique script in Japanese affects the neural processing of the complex word, the modality-independency for the complex words should be investigated for the other languages. Second, to further examine the similarity between words and phrases/sentences, we must investigate the complex words based on the morpheme-by-morpheme paradigm employed in experiment 3. As the experiment suggested the plausibility of the distributed morphology, which assumes words have syntactic structures, the following complex words must be investigated and compared to phrases/sentences with similar structures: morpheme-nested words (e.g., [un-[[predict]-able]]), and ambiguous words (e.g., unlockable: [un-[[lock]-able]] or [[un-[lock]]-able]). Moreover, complex words which lack their heads are interesting targets to investigate since phrases/sentences are always considered to have heads in their structures [e.g., exocentric compounds: pickpocket, “*neko-zyarasi*,” green foxtail (literally “cat teaser”); copulative compounds (dvandva): bittersweet, “*kusa-ki*,” grasses and trees]. As Japanese is an agglutinative language with a richer morphology than English, this language is adequate for the investigation. Moreover, based on the first direction, we must deal with

broader languages, such as the Ainu language, which is ambiguous between the boundary of words and sentences, is important since these languages extend the current results in language universality. Third, we must conduct further analysis for experiment 3 by applying analysis of ROIs to examine the difference in input between modalities. In experiment 3, the difference could not be investigated as we conducted the whole head analysis to demonstrate the modality-independency in language processing. On the other hand, the result of experiment 3 suggested that the auditory modality is a default modality for language processing. Therefore, the analysis will provide evidence of how different linguistic inputs are integrated and mapped into a similar neural activation. Fourth, to localize what brain region involves lexical processing, we must investigate whether the two temporal clusters in experiment 2 correspond to the two lexical processing stages. Due to the poor spatial resolution of EEG, we could only postulate that the temporal clusters may correspond to the two stages, lexical access and combination. Specifically, the dissertation's interpretation validity can be examined by applying MEG studies, such as Neophytou et al. 2019 and Stockall et al. 2019. A recent MEG study reported a common neural

activation for building compounds and phrases in the left temporal lobe (Flick et al., 2018; see also Gwilliams, 2020), although whether the left frontal language area activates during the morphological processing remained unclear. The result of experiment 2 replicated the studies and suggested a similar interpretation. Furthermore, previous fMRI studies demonstrated that the left inferior frontal gyrus is a core region for syntactic operation (Ohta et al., 2013a, 2013b; Ohta et al., 2017; Tanaka et al., 2017; Tanaka et al., 2019), suggesting that this region also plays an essential role in morphological processing. Therefore, in our future study, we must examine the function of the left inferior frontal gyrus in morphological processing.

References

- Albright, A., & Hayes B. 2003. Rules vs. analogy in English past tenses: A computational/experimental study. *Cognition*, 90(2), 119–161.
- Aronoff, M. (Ed.), 1976. *Word formation in generative grammar*. Cambridge, MA: MIT Press.
- Baayen, R. H. 2008. *Analyzing Linguistic Data: A Practical Introduction to Statistics Using R*. Cambridge University Press.
- Barr, D., J., Levy, R., Scheepers, C., & Tily, H. J. 2013. Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, 68, 255–278.
- Bates, D., & Mächler, M. 2009. lme4: Linear mixed-effects models using S4 classes [Software]. Available from <https://cran.r-project.org/web/packages/lme4/index.html>
- Bates, D, Mächler M., Bolker B., & Walker, S. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67. 1–48.
- Bates, D., Kliegl, R., Vasishth, S., & Baayen, H. 2018. Parsimonious mixed models. *arXiv preprint arXiv:1506.04967*.
- Buzsáki G, Draguhn A 2004. Neuronal oscillations in cortical networks. *Science*. 304(5679), 1926–1929.
- Butterworth, B. (Ed.), 1983. *Lexical representation. Language Production*. 257–294. Cambridge, MA: MIT Press.
- Chomsky, N. 1957. *Syntactic Structures*. Berlin: Mouton de Gruyter.
- Clahsen, H., & Ikemoto, Y. 2012. The mental representation of derived words: An experimental study of *-sa* and *-mi* nominals in Japanese. *The Mental Lexicon*, 7, 147–182.
- Cohen, J. 1992. A power primer. *Psychological Bulletin*, 112(1), 155–159.
- Coulmas, F. 2003. Mixed systems. In Coulmas, F. (Ed.), *Writing Systems* 168–189. Cambridge: Cambridge University Press.
- Crepaldi, D., Rastle, K., Coltheart, M., & Nickels, L. 2010. ‘Fell’ primes ‘fall’, but does ‘bell’ prime ‘ball’? Masked priming with irregularly-inflected primes. *Journal of Memory and Language*, 63, 83–99.
- Ding, N., Melloni, L., Zhang, H., Tian, X., & Poeppel, D. 2016. Cortical tracking of hierarchical linguistic structures in connected speech. *Nature Neuroscience*, 19, 158–164.
- Ding, N., Melloni, L., Yang, A., Wang, Y., Zhang, W., & Poeppel, D. 2017. Characterizing neural entrainment to hierarchical linguistic units using electroencephalography (EEG). *Frontiers in Human Neuroscience*, 11, 1–9.
- Embick, D., Ava C., and Amy J. Goodwin Davies. 2021. Morphology and the mental lexicon: Three questions about *decomposition*. In Anna Papafragou, John C. Trueswell and Lila R. Gleitman (eds.), *The Oxford handbook of the mental lexicon*, 77–97. Oxford: Oxford University Press.
- Embick, D., & Poeppel, D. 2015. Towards a computational(ist) neurobiology of language: Correlational, integrated, and explanatory neurolinguistics. *Language, Cognition and Neuroscience*, 30(4), 357–366.
- Flick, G., Oseki, Y., Kaczmarek, A., Al Kaabi, M., Marantz, A., & Pykkänen, L. (2018). Building words and phrases in the left temporal lobe. *Cortex*, 106, 213–236.
- Fiorentino, R., Naito-Billen, Y., & Minai, U. 2016. Morphological Decomposition in Japanese De-adjectival Nominals: Masked and Overt Priming Evidence. *Journal of Psycholinguistic Research*, 45, 575–597.
- Forster, K. I. & Davis, C. 1984. Repetition Priming and Frequency Attenuation in Lexical Access. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 10, 680–698.
- Frost, R., Forster, K. I., & Deutsch, A. 1997. What can we learn from the morphology of Hebrew? A masked-priming investigation of morphological representation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 23, 829–856.
- Fruchter, J., Stockall, L., & Marantz, A. 2013. MEG masked priming evidence for form-based decomposition of irregular verbs. *Frontiers in Human Neuroscience*, 7, 1–16.
- Fruchter, J., & Marantz, A. 2015. Decomposition, lookup, and recombination: MEG evidence for the full decomposition model of complex visual word recognition. *Brain and Language*, 143, 81–96.

- Giraud, A. L., & Poeppel, D. 2012. Cortical oscillations and speech processing: Emerging computational principles and operations. *Nature Neuroscience*, 15(4), 511–517.
- Gramfort, A., Luessi, M., Larson, E., Engemann, D. A., Strohmeier D., Brodbeck, C., Goj R., Jas, M., Brooks, T., Parkkonen, L., & Hämäläinen, M. 2013. MEG and EEG data analysis with MNE-Python. *Frontiers in Neuroscience*, 7, 1–13.
- Greenhouse, S. W., & Geisser, S. 1959. On methods in the analyses of profile data. *Psychometrika*, 24, 95–112.
- Gwilliams, L., Lewis, G., & Marantz, A. 2016. Functional characterization of letter-specific responses in time, space, and current polarity using magnetoencephalography. *Neuroimage*, 132, 320–333.
- Gwilliams, L., & Marantz, A. 2018. Morphological representations are extrapolated from morpho-syntactic rules. *Neuropsychologia*, 114, 77–87.
- Gwilliams, Laura. 2020. How the brain composes morphemes into meaning. *Philosophical Transactions of the Royal Society B: Biological Sciences*. 375:20190311.
- Hagiwara, H., Ito, T., Sugioka, Y., Kawamura, M. & Shiota, J. 1999. Neurolinguistic Evidence for Rule-Based Nominal Suffixation. *Language*, 75, 739–763.
- Hakala, T., Hulten, A., Lehtonen, M., Lagus, K., & Salmelin, R. 2018. Information properties of morphologically complex words modulate brain activity during word reading. *Human Brain Mapping*, 39, 2583–2595.
- Halle, M. & Marantz, A. 1993. Distributed Morphology and the Pieces of Inflection. In Hale, K. & Keyser S. (Eds.), *The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger*. 111–176. Cambridge, MA: MIT Press.
- Ito, T. (Ed.), 2002. *Bunporiron: rekisikon to tougo*. [Theory of grammar: lexicon and syntax]. Tokyo: Tokyodaigaku syuppankai.
- Ito, T. & Sugioka, Y. (Eds.). 2002. *Gono sikumi to gokeisei*. [The word structure and word formation]. Tokyo: Kenkyusha.
- Ito, T. & Sugioka, Y. 2016. “Go”ha atamanonakade dousyorisarerunoka. [How “words” are processed in the brain]. In Kageyama T. (Ed.), *Lexicon Forum* 7. 75–96. Tokyo: Hitsujishobo.
- Kageyama, T. (Ed.), 1993. *Bunpo to gokeisei* [Grammar and word formation]. Tokyo: Hitsujishobo.
- Kastner, I., Pyllkkänen, L., & Marantz, A. 2018. The form of morphemes: MEG evidence from masked priming of two Hebrew Templates. *Frontiers in Psychology*, 9, 1–17.
- Lau, E. F., Colin, P., & Poeppel, D. 2008. A cortical network for semantics:(de)constructing the N400. *Nature reviews neuroscience*, 9, 920–933.
- Lavric, A., Clapp, A., & Rastle, K. 2007. ERP Evidence of Morphological Analysis from Orthography: A Masked Priming Study. *Journal of Cognitive Neuroscience*, 19, 866–877.
- Leminen, A., Smolka E., Duñabeitia A. J. & Pliatsikas C. 2018. Morphological processing in the brain: The good (inflection), the bad (derivation) and the ugly (compounding). *Cortex*, 111, 4–44.
- Lewis, G., Solomyak O. & Marantz A. 2011. The neural basis of obligatory decomposition of suffixed words. *Brain and Language*, 118(3), 118–127.
- Lo, C., Tung, T., Ke, A. H., & Brennan, J. R. 2022. Hierarchy, Not Lexical Regularity, Modulates Low-Frequency Neural Synchrony During Language Comprehension. *Neurobiology of Language*, 3(4), 538–555.
- Luo, C., & Ding, N. 2020. Cortical encoding of acoustic and linguistic rhythms in spoken narratives. *eLife*, 9, Article e60433.
- Maekawa, Kikuo, Makoto Yamazaki, Toshinobu Ogiso, Takehiko Maruyama, Hideki Ogura, Wakako Kashino, Hanae Koiso, Masaya Yamaguchi, Makiro Tanaka and Yasuharu Den. 2014. Balanced corpus of contemporary written Japanese. *Language Resources and Evaluation* 48. 345–371.
- Maris, E. & Oostenveld, R. 2007. Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods*, 164, 177–190.
- Marslen-Wilson, W. 1993. Issues of process and representation in lexical access. In G. T. M. Altmann & R. Shillcock (Eds.), *Cognitive models of speech processing: The Second Sperlonga Meeting*. 187–210. Hove: Lawrence Erlbaum Associates Publishers.
- Marslen-Wilson, William and Tyler K. Lorraine 1998. Rules, representations, and the English past tense. *Trends in Cognitive Sciences*, 2(11), 428–435.

- Martin, A., E. 2020. A compositional neural architecture for language. *Journal of Cognitive Neuroscience*, 32(8), 1407–1427.
- Martin, A., E., & Doumas, L., A. 2017. A mechanism for the cortical computation of hierarchical linguistic structure. *PLOS Biology*, 15(3), Article e2000663.
- Melchert, H. C. (Eds.), 2003. *The Luwians*. Leiden: Brill.
- Meyer, L. 2018. The neural oscillations of speech processing and language comprehension: state of the art and emerging mechanisms. *European Journal of Neuroscience*, 48, 2609–2611.
- Meyer, L., Sun, Y., & Martin, A. E. 2020. Synchronous, but not entrained: Exogenous and endogenous cortical rhythms of speech and language processing. *Language, Cognition and Neuroscience*, 35(9), 1089–1099.
- Morris, J., Farnk, T., Grainger, J., & Holcomb, P. J. 2007. Semantic transparency and masked morphological priming: An ERP investigation. *Psychophysiology*, 44, 506–521.
- Morris, J. & Stockall, L. 2012. Early, equivalent ERP masked priming effects for regular and irregular morphology. *Brain and Language*, 123, 81–93.
- Nakajima, J. & Ohta, S., 2023. (Dis)similarities between Semantically Transparent and Lexicalized Nominal Suffixation in Japanese: An ERP Study Using a Masked Priming Paradigm. Masatoshi Koizumi (ed.), *Issues in Japanese Psycholinguistics: From Comparative Perspectives*. Mouton-NIJAL Library of Linguistics.
- Nakano, Y., Ikemoto, Y., Gunnar, J., & Clahsen, H. 2016. How orthography modulates morphological priming: Subliminal kanji activation in Japanese. *Frontiers in psychology*, 7, 1–10.
- Nakano, Y. & Kishimoto, K. 2019. Morpho-orthographic decomposition and kanji activation in visual processing of derived words in Japanese. Presented at *the 158th Meeting of the Linguistic Society of Japan*, Kunitachi-shi, Tokyo.
- National Institute for Japanese Language and Linguistics. 2018. The Balanced Corpus of Contemporary Written Japanese (BCCWJ): short unit word list ver. 1.1
- Neophytou, K., Manouilidou, C., Stockall, L., & Marantz, A. 2018. Syntactic and semantic restrictions on morphological recomposition: MEG evidence from Greek. *Brain and Language*, 183, 11–20.
- Ohta, S., Fukui, N. & Sakai K., L. 2013a. Computational principles of syntax in the regions specialized for language: Integrating theoretical linguistics and functional neuroimaging. *Frontiers in Behavioral Neuroscience*, 7: 204.
- Ohta, S., Fukui, N. & Sakai, K., L. 2013b. Syntactic computation in the human brain: The degree of merger as a key factor. *PLOS ONE*, 8: e56230.
- Ohta, S., Koizumi, M. & Sakai. K. L. 2017. Dissociating effects of scrambling and topicalization within the left frontal and temporal language areas: An fMRI study in Kaqchikel Maya. *Frontiers in Psychology*, 8: 748.
- Ohta, S., Oseai, Y., & Marantz, A. 2019. Dissociating the effects of morphemes and letters in visual word recognition: An MEG study of Japanese verbs. Presented at *the 25th meeting of Architectures and Mechanisms for Language Processing*, Moscow, Russia.
- Oldfield, R. 1971. The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*, 9, 812–815.
- Oseki, Y. & Tagawa T. 2019. Dual suppletion in Japanese. *The 14th Workshop on Altaic Formal Linguistics*, 193–204.
- Payne A. (Ed.), 2014. *Hieroglyphic Luwian: An Introduction with Original Texts*, Gaschwitz: Otto Harrassowitz.
- Peirce, J. W. 2007. PsychoPy-Psychophysics software in Python. *Journal of Neuroscience Methods*, 162, 8–13.
- Picton, T. W., Bentin, S., Berg, P., Donchin, E., Hillyard, S. A., Johnson, R., Miller, G. A., Ritter, W., Ruchkin, D. S., Rugg M. D., & Taylor, M., J. 2000. Guidelines for using human event-related potentials to study cognition: Recording standards and publication criteria. *Psychophysiology*, 37, 127–152.
- Pinker, S. 1991. Rules of language. *Science*, 253, 530–534.
- Pinker, S. & Prince, A. 1988. On language and connectionism: analysis of a parallel distributed processing model of language acquisition. *Cognition*, 28, 73–193.

- Pinker, S. & Ullman M. T. 2002. The past and future of the past tense. *Trends in Cognitive Sciences*, 6(11), 456–463.
- Poeppel, D. 2012. The maps problem and the mapping problem: Two challenges for a cognitive neuroscience of speech and language. *Cognitive Neuropsychology*, 29(1–2), 34–55.
- Poeppel, D., & Embick, D. 2005. Defining the relation between linguistics and neuroscience. In A. Cutler (Ed.), *Twenty-first century psycholinguistics: Four cornerstones*. 103–120. Routledge.
- Pylkkänen, L. & Marantz, A. 2003. Tracking the time course of word recognition with MEG. *Trends in cognitive sciences*, 7, 187–189.
- Rastle, K., Davis M. H., & New, B. 2004. The broth in my brother's brothel: Morpho-orthographic segmentation in visual word recognition. *Psychonomic Bulletin and Review*, 11(6), 1090–1098.
- Rastle, K., Davis, M., H., Marslen-Wilson, D., & Tyler, L. K. 2000. Morphological and semantic effects in visual word recognition: A time course study. *Language and Cognitive Processes*, 15, 507–537.
- Shaffer, J. P. 1986. Modified sequentially rejective multiple test procedures. *Journal of the American Statistical Association*, 81, 826–831.
- Shibatani, M. 1990. *The Languages of Japan*. Cambridge: CUP.
- Solomyak, O. & Marantz, A. 2009. Lexical access in early stages of visual word processing: A single-trial correlational MEG study of heteronym recognition. *Brain and Language*, 108, 191–196.
- Solomyak, O. & Marantz, A. 2010. MEG evidence for early morphological decomposition in visual word recognition: A single-trial correlational MEG study. *Journal of Cognitive Neuroscience*, 22, 2042–2057.
- Stockall, L., Manouilidou, C., Gwilliams, L., Neophytou, K., & Marantz, A. 2019. Prefix Stripping Re-Re-Revisited: MEG Investigations of Morphological Decomposition and Recomposition. *Frontiers in Psychology*, 10, 1–17.
- Stockall, L. & Marantz, A. 2006. A single route, full decomposition model of morphological complexity: MEG evidence. *The Mental Lexicon*, 1, 85–123.
- Sugioka, Y. (Eds.), 1986. *Interaction of Derivational Morphology and Syntax in Japanese and English*. New York: Garland Publishing.
- Taft, M. & Forster, K. I. 1975. Lexical storage and retrieval of polymorphemic and polysyllabic words. *Journal of Verbal Learning and Verbal Behavior*, 14, 638–647.
- Tagawa, T. 2008. *Bunsankeitaironniyoru doushino katsuyouto gokeiseionokennkyuu*. [Study of inflection of verbs and word formation by distributed morphology]. 90–119. Doctoral Dissertation, University of Tsukuba.
- Tanaka, Kyohei, Issa Nakamura, Shinri Ohta, Naoki Fukui, Mihoko Zushi, Hiroki Narita & Kuniyoshi L. Sakai. 2019. Merge-generability as the key concept of human language: Evidence from neuroscience. *Frontiers in Psychology*, 10: 2673.
- Tanaka, Kyohei, Shinri Ohta, Ryuta Kinno & Kuniyoshi L. Sakai. 2017. Activation changes of the left inferior frontal gyrus for the factors of construction and scrambling in a sentence. *Proceedings of the Japan Academy, Series B*, 93. 511–522.
- Wray, S., Stockall, L., & Marantz, A. 2022. Early Form Based Morphological Decomposition in Tagalog: MEG evidence from Reduplication, Infixation, and Circumfixation. *Neurobiology of Language*. *Neurobiology of Language*, 3, 235–275.
- Vartiainen, J., Aggajaro, S., Lehtonen, M., Hultén, A., Laine, M., & Salmelin, R., 2009. Neural dynamics of reading morphologically complex words. *Neuroimage*, 47, 2064–2072.
- Voeten, C. C. 2021. *permutes: Permutation Tests for Time Series Data* [Software]. Available from <https://cran.r-project.org/package=permutes> (last modified 15 June 2022)
- Zweig, E. & Pylkkänen, L. 2009. A visual M170 effect of morphological complexity. *Language and Cognitive Processes*, 24, 412–439.

Acknowledgements

I would like to thank my supervisor, Associate Professor Shinri Ohta, for his suggestive and insightful advice through the years. His patience and forbearance attitude has been invaluable to me. I would like to thank the committee of my dissertation: Professor Ayumi Ueyama, Associate Professor Michinori Shimoji, Associate Professor Masako Maeda (English Language and Literature, Kyushu University), and Lecturer Yohei Oseki (Language Computational Cognitive Science Laboratory, University of Tokyo). Professor Ayumi Ueyama is the very first person who led me to the world of linguistics and is a very generous mentor. I owe much to her advice, especially for research direction. Associate Professor Michinori Shimoji showed me how researchers should be through his attitude to his students. I also thank Assistant Professor Emi Yamada for her beneficial suggestions, especially for EEG analyses. In addition, I would like to thank my dissertation committee.

I have greatly benefitted from Dr. Chia-Wen Lo (The University of Michigan), Dr. Yuko Urabe (Tokyo University of Foreign Studies), Associate Professor Toshikazu Ikuta (The University of Mississippi), Dr. Daniel Cory Gallagher, Dr. Yuki Hirose (The University of Tokyo), Dr. Takumi Tagawa (The University of Tsukuba) and Mr. Koki Yamaguchi for their invaluable

comments on my research. I would also like to thank the members of my laboratory. Each of their research was motivational and fascinating.

I would like to offer the same gratitude to members of my family who passed away during my graduate course. I hope my only relative, Mizuki, my younger brother, lives a delightful life.

In my personal life, I would like to thank Mr. Yudai Tagawa Matsusaki, Dr. Miki Ogasawara, Mrs. Miki Ida, Mrs. Yuko Saito, Mr. Yusuke Sato, Mrs. Risako Mukaida (Sakai), Ms. Nozomi Kodama, Ms. Akie Koga, Mrs. Megan Eiko Uchida, and Mr. Yusuke Terashima for their encouragements. Moreover, many anonymous people encouraged me to get through tough situations. Last but not least, Mr. Yuki Shishido and Ms. Yuko Fujita enlightened and helped me in getting through these tough, troubled times, before and during my graduate course, and I owe them a lot.

Finally, I would like to quote a passage from “*Propos sur le bonheur*.” This passage is quoted from my very first textbook of Latin, which led me to the world of linguistics. Its content has been one of my mottos since I do sympathize with this passage, and I have been far from “*le bonheur*” for the last decade. Still far away from “*le bonheur*,” I am, but someday I hope I am. Je fais mon bonheur moi-même. C’est un mon “malgrava peto.”

“On dit que le bonheur nous fuit toujours. Cela est vrai du bonheur reçu, parce qu’il n’y a point de bonheur reçu. Mais le bonheur que l’on se fait ne trompe point. C’est apprendre, et l’on apprend toujours. Plus on sait, et plus on est capable d’apprendre.” (*Propos sur le bonheur*, “Aristote,” 15th September 1924, Alain).

Appendix Juni

FVI HOC, SED NON SUM. ET, NEC ERO.

SPES PRODIT. ITA FRAGOSA ET BREVI MENTE VIAVM, TVNC MORIAR.

Est videor insolitum scribere appendiculam confessionis in Latin. Scripsi maledictia in hac sectione. Tamen nolo aliquis legere quod scripsi in hac sectione cum plenes maledicta.

Maledicta sunt qui habeo multa odium per circumcisona et universita Ciusiu.

Per hos duo et tres anni, timui ut deficere animo fracto. Aut animus hic timere a deficientes. Quoniam meus prior circumiecta fuit pessima maxime alienus et factum sum vulneratis, est perterrentes facere novum inceptm mihi. Haec cura fuit rectus cupide. Metus deficere toleravi. Seneca inquit “spem metus sequitur”. Vere hababam spem et metum. Sed spes prodit.

Odi. Ego odi quos consulto nocerunt vitae mei. MALE SINT VOBIS. LONGA ET GRAVA TORMENTA, IANO, M., KONDO, T., et ITO, T., SINT IN VITIS VOBIS.

IANO M. qui nocerunt et nocet me obstanto experimentis identidem. Praeterea hic homines calumniatus est me in twittere. Nec corpus querulo nec mens vacet aegra dolore noxque die gravior sit tibi, nocte dies (Ovid. Ibis 115-116). KONDO T. est primus homo qui destruxit vitae mei et calumniatus est me identidem. Sisque miser semper nec sis miserabilis ulli: gaudet adversis femina virque tuis (Ovid. Ibis 117-118). Et ITO T. est homo qui identidem calumniatus est me in twittere. Accedat lacrimis odium, dignusque putere qui, mala cum tuleris plurima, plura feras. sitque, quod est rarum, solito defecta favore Fortunae facies invidiosa tuae. (Ovid. Ibis 119-120). Causaque non desint, desint vobis copiae mortis: optatam fugiant vitae coacta necem. luctatusque diu cruciatos spiritus artus deserat et longa torqueat ante mora. Non procul a proprio stipite poma cadunt. Hi mali hominis qui consulto nocerunt me fuisse calamitatibus cum vestrum socii.

Et odi mei universitam universita qui mihi est non gymnasium per investigare nec erudire

sed est negligens fraudator.

Utinam ne vere scriberem in hac sectione. Sed non. Eram et sum infelix et qui hominis mali infelix truserunt.

Postremo, beate sit mihi. Est mens mei nunc fracta a novem anni horribiles et sentiam, ut, si omina fuisset, procul, procul a me. Si mei vita darabit, moriar.

Appendices

Table S1. Full list of the stimuli used in experiment 1.

Derived complex end with “sa”	Derived complex end with “mi”	Non-derived words end with “sa”	Non-derived words end with “mi”
Aka-sa (あかさ, red)	Aka-mi (あかみ, reddish)	Asa (あさ, morning)	Ami (あみ, rete)
Akaru-sa (あかるさ, bright)	Akaru-mi (あかるみ, become known)	Esa (えさ, bait)	Gomi (ごみ, waste)
Ama-sa (あまさ, sweetness)	Ama-mi (あまみ, sugariness)	Hana-busa (はなぶさ, flower cluster)	Hasami (はさみ, scissors)
Ao-sa (あおさ, blueness)	Ao-mi (あおみ, blueness/greenness)	Hayabusa (はやぶさ, falcon)	Hitomi (ひとみ, eye)
Atataka-sa (あたたかさ, warmth)	Atataka-mi (あたたかみ, kindness)	Higasa (ひがさ, sunshade)	Izumi (いずみ, spring)
Atu-sa (あつさ, thickness)	Atu-mi (あつみ, thickness/profundity)	Hosikusa (ほしくさ, hay)	Kagami (かがみ, mirror)
Huka-sa (ふかさ, depth)	Huka-mi (ふかみ, depth/deep place)	Husa (ふさ, bunch)	Kami (かみ, paper)
Ita-sa (いたさ, painfulness)	Ita-mi (いたみ, oain)	Ikusa (いくさ, war)	Kami-gami (かみがみ, gods)
Kanasi-sa (かなしさ, sadness)	Kanasi-mi (かなしみ, gloominess)	Kare-kusa (かれくさ, dead grass)	Kata-sumi (かたすみ, corner)
Kara-sa (からい, hot taste)	Kara-mi (からみ, pungent taste)	Kasa (かさ, umbrella)	Kimi (きみ, you)
Kayu-sa (かゆさ, itch)	Kayu-mi (かゆみ, itching sensation)	Kesa (けさ, this morning)	Kura-yami (くらやみ, dusk)
Kurusi-sa (くるしさ, pain)	Kurusi-mi (くるしみ, suffering)	Kuma-zasa (くまざさ, Veitch's bamboo)	Mimi (みみ, ear)
Kusa-sa (くささ, smell)	Kusa-mi (くさみ, unpleasant smell)	Kusa (くさ, grass)	Minami (みなみ, south)
Maru-sa (まるさ, roundness)	Maru-mi (まるみ, roundish)	Mai-asa (まいあさ, every morning)	Mizuumi (みずうみ, lake)
Niga-sa (にがさ, bitterness)	Niga-mi (にがみ, bitter taste)	Miti-kusa (みちくさ, dawdle)	Nami (なみ, wave)

Omo-sa (おもさ, weight)	Oso-mi (おもみ, weight/importance)	Mizu-kusa (みずくさ, aquatic plant)	Nezumi (ねずみ, mouse)
Omosiro-sa (おもしろさ, interest)	Omosiro-mi (おもしろみ, attraction/fun)	Nagisa (なぎさ, beach)	Ookami (おおかみ, wolf)
Sibu-sa (しぶさ, astringency)	Sibu-mi (しぶみ, tartness)	Nana-kusa (ななくさ, seven spring herbs)	Sumi (すみ, corner)
Sitasi-sa (したしさ, closeness)	Sitasi-mi (したしみ, intimacy)	Sasa (ささ, bamboo grass)	Tatami (たたみ, tatami mat)
Sugo-sa (すごさ, terribleness)	Sugo-mi (すごみ, ghastliness)	Sigusa (しぐさ, behavior)	Te-gami (てがみ, letter)
Taka-sa (たかさ, height)	Taka-mi (たかみ, high place)	Tibusa (ちぶさ, breast)	Tomi (とみ, wealth)
Tanosi-sa (たのしさ, enjoyment)	Tanosi-mi (たのしみ, recreation)	Tubasa (つばさ, wing)	Tubomi (つぼみ, bud)
Tsuyo-sa (つよさ, strength)	Tuyo-mi (つよみ, advantage)	Tukasa (つかさ, leader)	Tumi (つみ, sin)
Uma-sa (うまさ, deliciousness)	Uma-mi (うまみ, good taste)	Uwasa (うわさ, rumor)	Umi (うみ, sea)
Yowa-sa (よわさ, weakness)	Yowa-mi (よわみ, weak point)	Yoku-asa (よくあさ, next morning)	Yami (やみ, dark)

Table S2. Full list of the adjectives used in experiment 2.

-mi type	Non-mi type stem with a simple stem	Non-mi type stem with a complex stem
Aka-i (赤い, red)	Atarasi-i (新しい, new)	Ao-ziro-i (青白い, pale)
Akaru-i (明るい, bright)	Haya-i (早い, early)	Ara-arasi-i (荒々しい, violent)
Ama-i (甘い, sweet)	Hido-i (酷い, horrible)	Baka-bakasi-i (馬鹿馬鹿しい, ridiculous)
Ao-i (青い, blue)	Hiku-i (低い, low)	Bu-atu-i (分厚い, thick)
Atataka-i (温かい, warm)	Hiro-i (広い, spacious)	Haba-hiro-i (幅広い, broad)
Atu-i (厚い, thick)	Huru-i (古い, old)	Hada-zamu-i (肌寒い, chilly)
Huka-i (深い, deep)	Kata-i (固い, hard)	Hodo-too-i (程遠い, far away)
Ita-i (痛い, painful)	Kawai-i (かわいい, cute)	Hoso-naga-i (細長い, long and thin)
Kanasi-i (悲しい, sad)	Kowa-i (怖い, scared)	Kan-daka-i (甲高い, high-pitched)
Kara-i (からい, spicy)	Kura-i (暗い, dark)	Ki-iro-i (黄色い, yellow)
Kayu-i (痒い, itchy)	Mizika-i (短い, short)	Ki-mazu-i (気まずい, unpleasant)
Kurusi-i (くるしい, distressed)	Muzukasi-i (むずかしい, difficult)	Ko-daka-i (小高い, slightly elevated)
Kusa-i (臭い, smelly)	Naga-i (長い, long)	Kokoro-boso-i (心細い, lonely)
Maru-i (丸い, round)	Oisi-i (おいしい, tasty)	Mono-sugo-i (物凄い, tremendous)
Niga-i (にがい, bitter)	Oo-i (多い, many/much)	Mu-siatu-i (蒸し暑い, hot and humid)
Okasi-i (おかしい, funny)	Ooki-i (大きい, large)	Na-daka-i (名高い, renowned)
Omo-i (重い, heavy)	Oso-i (遅い, late/slow)	Nama-namasi-i (生々しい, graphic/vivid)
Omosiro-i (おもしろい, interesting)	Sukuna-i (少ない, few/little)	Ne-zuyo-i (根強い, deep-rooted)
Sibu-i (渋い, astringent)	Tiisa-i (小さい, small)	Oku-buka-i (奥深い, profound)
Sitasi-i (親しい, friendly)	Too-i (遠い, distant)	Si-kaku-i (四角い, square)
Sugo-i (凄い, fantastic)	Tuta-i (つらい, painful)	Su-baya-i (素早い, quick)
Taka-i (高い, high)	Uresi-i (嬉しい, happy)	Te-atu-i (手厚い, hospitable)

Toro-i (とろい, stupid/slow)	Usu-i (薄い, thin)	Te-baya-i (手早い, speedy)
Tuyo-i (強い, strong)	Waka-i (若い, young)	Tya-iro-i (茶色い, brown)
Uma-i (旨い, tasty)	Waru-i (わるい, bad)	Usu-gura-i (薄暗い, dim)
Yowa-i (弱い, weak)	Yasu-i (安い, cheap)	Yowa-yowasi-i (弱々しい, faint)

Table S3. Full list of roots of the stimuli used in experiment 3.

Abi-ru (浴びる, shower)	Oe-ru (終える, finish)
Age-ru (上げる, raise)	Sage-ru (下げる, turn down)
Ake-ru (開ける, open)	Sake-ru (避ける, avoid)
Ate-ru (当てる, hit)	Seme-ru (攻める, attack)
Hime-ru (秘める, hide)	Sime-ru (占める, occupy)
Home-ru (褒める, praise)	Soe-ru (添える, attach)
Hure-ru (触れる, touch)	Some-ru (染める, dye)
Huse-ru (伏せる, conceal)	Sue-ru (据える, install)
Ire-ru (入れる, put in)	Sute-ru (捨てる, throw away)
Kae-ru (変える, change)	Taberu (食べる, eat)
Kake-ru (掛ける, hang)	Tae-ru (耐える, bear)
Kari-ru (借りる, rent)	Tame-ru (溜める, store)
Kime-ru (決める, decide)	Tate-ru (立てる, stand)
Kise-ru (着せる, cloth)	Toge-ru (遂げる, finish)
Koeru (越える, pass over)	Toji-ru (閉じる, close)
Kome-ru (込める, charge)	Tuge-ru (遂げる, achieve)
Mage-ru (曲げる, bend)	Tuke-ru (付ける, attach)
Maze-ru (混ぜる, mix)	Ture-ru (連れる, bring)
Mise-ru (見せる, show)	Uke-ru (受ける, receive)
Muke-ru (向ける, point)	Ume-ru (埋める, bury)
Nade-ru (撫でる, stroke)	Wabi-ru (詫びる, apologize)
Nage-ru (投げる, throw)	Wake-ru (分ける, separate)
Nobe-ru (述べる, state)	Yose-ru (寄せる, gather)
Nose-ru (乗せる, load on)	Yude-ru (茹でる, boil)