문학박사 학위논문

The embodiment cognition of abstract concepts in L2 processing

– Perceptual strength norms and EEG analysis on modality specific neural responses –

제2언어처리에서의 추상 개념의 체화된 인지: 단어 감각 지수와 감각 특정 신경반응에 대한 EEG 분석

2023년 2월

서울대학교 대학원

영어영문학과 어학전공

이 종 현

The embodiment cognition of abstract concepts in L2 processing

- Perceptual strength norms and EEG analysis on modality specific neural responses -

지도 교수 송 미 정

이 논문을 문학박사 학위논문으로 제출함
2022년 10월

서울대학교 대학원
영어영문학과 어학 전공
이 종 현

이종현의 문학박사 학위논문을 인준함
2023년 2월

위 원 장 ___이 성 은___ (인)

부위원장 ___이 정 미___ (인)

위 원 ___신 정 아___ (인)

위 원 ___권 영 성___ (인)

위 원 ___송 미 정___ (인)

Abstract

This dissertation explores the embodiment of abstract concepts in Second Language (L2) processing through the utilization of perceptual strength norms and EEG experiments on modality-specific neural responses. According to the embodied cognition theory, language comprehension is achieved through the activation of the same neural regions that process perception, movement, and emotion, rather than solely through language-specific regions. This assumption is supported by studies demonstrating that the associated sensory-motor areas are activated during language processing. However, previous research has been limited to motor-action related words and has not been conducted with different types of language users such as L2 speakers. This dissertation aims to address these limitations by investigating the embodiment of L2 English with an emphasis on abstract concepts that have yet to be fully explored within the framework of embodiment cognition theory.

This dissertation first investigates whether abstract concepts are rooted in sensory-motor regions, similar to concrete concepts. Abstract concepts have been deemed as incapable of being comprehended through sensorimotor experiences, as they do not have a tangible referent. Several studies have recently attempted to explain abstract concepts through the activation of sensory-motor regions including an additional sense, interoception, which refers to the perception of internal bodily sensations. This study examines this issue by assessing perceptual strength norms that incorporate interoception in Experiment 1, as well as comparing modality-specific neural responses between audio and interoception in Experiment 2.

This study also explores whether L2 processing is embodied in sensory-motor regions in the same way as L1 processing. The participants in this study were L2 learners who learned English in a foreign environment, which presents a qualitative and quantitative difference in the language learning and usage environment compared to L1 learners. L2 learners have less direct contact with the environment, often mediated through their native language. Given that the embodiment of language is presumed to be shaped through repeated associations between language and environment, these differences in the environment may have a significant impact on the embodiment of L2 processing. For this purpose, the current study compares the distribution of perceptual strengths of L2 learners and native speakers in Experiment 1 and examines differences in modality-specific neural responses in Experiment 2.

Two experiments were conducted to explore these research questions. The objective of Experiment 1 was to construct perceptual strength norms of L1 Korean, L1 English, and L2 English and compare each of them. The perceptual strength norm is a sensory profile of a word measuring how words are experienced through the five common senses and interoception. It can be used as a resource for future psycholinguistic research and allows the investigation of the degree to which L1 and L2 language users associate language with human perceptual experiences including interoception. Experiment 2 aimed to examine whether modality-specific neural responses differed between audio and interoception words and how those responses differed between L1 and L2. To this end, an EEG experiment was conducted in which Korean native speakers performed a lexical decision task involving audio and interoception words in Korean and L2 English. According to the embodiment cognition theory, distinct neural responses will be observed between audio and interoception as audio being processed in audio-related regions and interoception in interoception-related regions.

The results of Experiments 1 and 2 revealed that abstract concepts were also grounded in sensory-motor regions. Experiment 1 found a significant correlation between interoception and abstractness, implying that interoception was an important factor in the embodiment of abstract concepts. Experiment 2 further supported the result of Experiment 1. Interoception words elicited distinct neural responses from audio, and its neural generators were estimated as interoception-related regions. These findings demonstrated that the embodiment of abstract concepts can be explained through the activation of sensory-motor areas including interoception.

There were no notable differences between L1 and L2 processing in perceptual strength norms and modality-specific responses. This suggests that L2 processing is embodied in a similar manner to L1 processing despite fewer environmental exposures. However, the experiments also demonstrate that the embodiment of L2 processing is not entirely identical to L1. In Experiment 1, the perceptual experiences of L2 learners were influenced by their native language, resulting in a different distribution from L1 English in several words. In Experiment 2, L2 processing exhibited differences in the topography and polarity of gamma band power in time-frequency analysis and additional activation of the left temporal pole, the region related to language processing, in source estimation analysis. The results suggest that the embodiment of L2 processing may be attenuated or relying more on linguistic experience in terms of embodiment.

In conclusion, the present dissertation aimed to investigate how the embodiment of L2 processing differed from L1 processing, with a focus on

interoception and its relationship with abstract concepts. Through the conduct of two experiments, first, it was determined that abstract concepts, similar to concrete concepts, are rooted in sensory-motor areas, and interoception plays a crucial role in the embodiment of abstract concepts. In addition, the findings of this dissertation indicated that L2 processing was embodied in a similar manner to L1 processing, but it may be characterized by weaker embodiment and a greater reliance on linguistic experience.

Keyword : embodiment cognition, abstract concepts, second language processing, EEG, perceptual strength norms, source estimation

Student Number : 2017-32674

Table of Contents

Chapter 1. Introduction

1.1. Goal

This dissertation aims to explore whether linguistic concepts including both concrete and abstract ones are "embodied" or "grounded" in modality-specific brain systems in second language (L2) processing[1] and how it is different from first language (L1) processing. More specifically, it investigates (i) whether the embodiment of lexical processing is essential for comprehension, (ii) whether abstract concepts are embodied as well as concrete concepts and if so, on what they are grounded, and (iii) whether L2 processing is embodied as L1 processing and if so, how it differs from L1 processing. To accomplish these objectives, the study will (i) classify concrete and abstract concepts according to their sensory modality and establish multilingual perceptual strength norms, and (ii) utilizing these norms, examine the neural representation of these concepts, with a focus on audio-related words and interoception-related words, in sensory-motor regions of the brain, comparing L1 Korean and L2 English of Korean native speakers.

---

[1] In the Second Language Acquisition (SLA) literature, the terms, second language (L2) and foreign language, are often distinguished. In some contexts, second language refers to a language that is learned by non-native speakers in the places where that language is spoken as the official language, whereas foreign language is defined as the one learned and used in foreign situations where the language is not spoken as a native language. According to this distinction, the term L2 used here is inappropriate in some respects. This study is conducted in the environment of learning L2 as a foreign language, mainly addressing the issues of L2 learners who learn L2 in foreign situations. Thus, the term L2 used here is actually closer to the definition of a foreign language, and it might be more appropriate to refer to L2 as a foreign language, instead. However, this dissertation used the term L2 instead of a foreign language, in order to more clearly refer to the language with the highest proficiency other than the native language, contrasting with L1 or other foreign languages. Also, it was considered that the terms are used without distinction in embodiment cognition research. Therefore, it is important to note that even if the findings are discussed with the term L2, it is more about foreign language in a strict sense. Although it does not mean that the findings in this study cannot be extended to L2 in a broader sense, including all learned languages besides one's native languages, the primary interpretation should be applied to a foreign language in a narrow sense.

The embodiment theory of cognition posits that language comprehension relies on neural mechanisms that are typically employed for perception, action, and emotion (Glenberg et al., 2008). While a growing body of research from behavioral and neuropsychological studies has established the importance of such mechanisms for L1 processing, particularly for concrete concepts, the extent to which they are engaged in L2 processing remains relatively unexplored. The learning environment of L2 adult learners differs significantly from that of L1 child learners in terms of the context of language exposure, the quantity and quality of input, and other factors. Given that the repetitive association between language and the experiences surrounding its use is a critical factor in explaining the embodiment of language (Matheson & Barsalou, 2018), the differences in the learning environment between L1 and L2 may contribute to variations in the embodiment of language comprehension. In this regard, investigating embodiment in L2 processing and comparing it to L1 processing will expand our understanding of L2 processing and provide insight into the neural mechanisms underlying language comprehension.

This study also aims to investigate the embodiment of L1 and L2 processing, with a specific focus on the processing of abstract concepts. While a number of studies have examined embodiment in L2 processing, there is a dearth of research that specifically addresses abstract concepts. Abstract concepts, which are broadly defined as those lacking a perceivable referent (Kiefer & Harpaintner, 2020), constitute a substantial portion of language, yet their embodiment remains under-explored in the literature. The presence of abstract concepts is often viewed as challenging the assumption that language comprehension is rooted in perception and action. Thus, providing evidence for the embodiment of abstract concepts in the brain is crucial for the advancement of embodiment cognition theory. The current study will address this issue through a series of behavioral and neural experiments, assessing the embodiment of abstract concepts in L1 and L2 processing.

1.2. Contribution

1.2.1. Theoretical contribution

First of all, this study makes a significant contribution to the field by addressing the under-explored area of abstract concepts in embodiment cognition theory. While a number of studies have examined the grounding of concrete concepts through behavioral and neuropsychological experiments, empirical evidence for abstract concepts in L1 is scarce. Additionally, there are a variety of conflicting theories on abstract concepts, such as *the Dual Code Theory, Context availability model, Affective Embodiment Account,* and *Word as Social Tool theory.* By investigating the embodiment of abstract concepts, this study aims to enhance our understanding of this topic in the field of embodiment cognition theory and to provide a basis for evaluating the relative compatibility of existing theories with empirical data.

This study also expands the examination of embodiment cognition to include L2 processing, which is both informative for understanding L2 processing and for the advancement of embodiment cognition theory. While L1 and L2 share some common properties of language processing, many aspects of them are qualitatively or quantitatively distinct. Identifying the sources of these differences is a crucial goal of L2 research. The current study will help to clarify the distinctions between L1 and L2 in terms of embodiment cognition, while also providing a deeper understanding of embodiment cognition theory by comparing L1 and L2 processing. As there are variations between L1 and L2 in terms of learning environment, timing of acquisition, and proficiency level, investigating the impact of these variables on linguistic embodiment may expand our knowledge of embodiment cognition.

1.2.2. Pedagogical contribution

This study provides pedagogical implications of embodiment cognition in L2 education. There have been various discussions about the underlying reasons for differences in performance between L1 and L2 speakers. The present study adds aspects of embodiment to such a discussion, exploring whether the degree to which language comprehension is embodied is the source of differences between L1 and L2. If difference in embodiment is a decisive factor in discrepancies between L1 and L2 processing, a teaching approach should take the embodiment into account. For example, it can be argued that a pedagogical approach that emphasizes the experience surrounding language, not simply mediated through native language, is required. In fact, there have been studies that regarded gesture as an important element of learning and applied it to the teaching method, based on embodiment cognition theory (Kelly et al., 2009). However, these studies tend to exploit the concept of embodiment superficially, equating gesture to embodiment processing, without a sound theoretical foundation. The present study will contribute to establishing a theoretical foundation for pedagogical methods by conducting a thorough examination of L2 embodied processing.

1.2.3. Methodological contribution

The present study aims to contribute to the advancement of EEG source analysis by examining the involvement of various brain regions during language processing. While EEG has traditionally been known for its high temporal resolution, it has been considered less accurate in terms of spatial resolution in comparison to other neuroimaging techniques such as fMRI, MEG, and PET. However, with the development of source analysis methods, EEG has increasingly been utilized in research on embodiment cognition. For example, Dalla Volta et al., (2014) employed source analysis to investigate the activation of specific brain regions in the processing of concrete and abstract concepts. This study seeks to expand upon such research by utilizing a combination of behavioral experimentation and various source analysis techniques to evaluate the consistency and validity of the findings.

1.2.3. Outline of the dissertation

In the following Chapter 2, the theoretical background on the embodiment cognition theory and related theories in second language acquisition (SLA) literature are described. In this chapter, various issues of embodiment cognition theory are introduced and the problems to be solved for the development of the theory are presented. This also examines how to incorporate embodiment cognition theory from the perspective of SLA by exploring several relevant theories. In Chapter 3, the neurological background for the EEG experiment (Experiment 2) is reviewed. Chapter 4 presents the method and results of Experiment 1 and examines its implications for embodiment processing. In Chapter 5, based on the results of Experiment 1, the EEG experiment is conducted, and its results are described and discussed. Chapter 6 comprehensively discusses the results of Experiment 1 and Experiment 2, and suggests the implications, limitations, and future research directions of this dissertation.

Chapter 2. Literature review

2.1. Embodiment cognition theory

2.1.1. Two perspectives on language representation

Imagine you are in a room trying to figure out what a Chinese word, “yǐzi” means. You don't know Chinese and maybe you don't even know it's Chinese. Still, at least if you know that it is a word or some kind of symbol, you may look up a Chinese dictionary which happens to be placed in the room and end up managing to find the “meaning” from it. But unfortunately, what you run into is other arbitrary symbols. You may be able to find again the newly encountered symbols in other pages of the dictionary, but there only other new symbols welcome you. Then, how is it possible to understand what “yǐzi” means in this situation or any circumstance? Should this arbitrary symbol be grounded on something in the world? This is the “Chinese room” argument by Harnad (1990) (Glenberg & Kaschak, 2002; an adapted version of Searle (1980)), which typically exemplifies “The symbol grounding problem” (Kaschak et al., 2014).

It has been held that linguistic meaning is manifested in our brain as abstract representations in language-specific modules such as Broca's area and Wernicke's area (Pulvermüller, 2005). According to this dominant perspective, language is primarily processed in domain-specific regions, and does not require significant involvement from the sensory-motor system. The meaning of language is conveyed through amodal and arbitrary symbols, such as words, which are arranged according to syntactic rules (Glenberg & Kaschak, 2002). This view is consistent with cognitivism, which emerged as a dominant approach in the mid-

20th century following criticisms of behaviorism by Chomsky in 1959. From this perspective, the mind is seen as a computational system in which computations on amodal representations, which are dissociated from the input of the sensory-motor system, give rise to cognition ("the modularity of the mind"; Fodor, 1983). Additionally, it is believed that a unique human language faculty serves as a computational system that internally produces amodal symbols and maps them to the sensory-motor system, which is separate from it (Hauser et al., 2002).

However, considering the analogy of the Chinese room, it seems difficult to understand the meaning of language completely independently of the world and excluding its connection with it. Arbitrary symbols cannot be understood exclusively through their connection to other arbitrary symbols within amodal computational system (Kaschak et al., 2014), as illustrated in the analogy above. But, what if, for the comprehension of "yǐzi", not another arbitrary symbol, but a pictogram symbolizing it was given? Then, it should be much easier to understand it. It is because a pictogram is grounded in our world knowledge to some extent although it is also arbitrary.

In contrast to the amodal view of language and cognition, the embodiment theory of cognition suggests that language processing should be understood in the sense of mind-body-world interaction. Those arbitrary symbols come to be meaningful when grounded in our sensory-motor knowledge about the world (Kaschak et al., 2014). For example, one understands the concept, HAMMER, by retrieving all the sensory and motor information he or she has about hammers: visual information of a hammer and nail, auditory information of banging of hammering against a wall, tactile information of heaviness and bluntness, motor information of grasping a hammer and swinging it with an arm, and painful feeling of hurting a finger with it. According to embodiment cognition theory, all this sensory-motor information, grounded in the world, is essential to comprehend a word, while in the disembodied view, HAMMER will be merely a symbolic representation (Mahon & Caramazza, 2008). In this regard, the embodiment cognition approach maintains that language comprehension is neuro-anatomically grounded in the sensory-motor system, not solely computed in language-specific regions of the brain. This approach holds that the same areas of the brain that are used for perception, action, and emotion are also activated during language processing. A number of experimental studies have provided evidence for this viewpoint by demonstrating the involvement of these regions during language processing (Bergen et al., 2010; Cervetto et al., 2021; Chambers et al., , 2004; Dalla Volta et al., 2014; Glenberg & Kaschak, 2002; Glenberg et al, 2008; Hauk et al., 2004; Hauk & Pulvermüller, 2004; Klepp et al., 2015; Pulvermüller et al., 2005; Pulvermüller et al., 2009; Tettamanti et al., 2005; Vukovic & Shtyrov, 2014).

A significant body of evidence for the embodiment theory of cognition comes from the findings that the motor cortex is automatically engaged when individuals perceive language stimuli that indicate bodily action, such as action verbs. This motor involvement has been observed in both behavioral and neuroimaging studies that have utilized various techniques such as fMRI, MEG, and EEG (Cervetto et al., 2021; Dalla Volta et al., 2014; Hauk et al., 2004; Hauk & Pulvermüller, 2004; Klepp et al, 2015; Pulvermüller et al., 2009; Tettamanti et al., 2005; Vukovic & Shtyrov, 2014; Zhang et al., 2020). For example, in an EEG study, it was found that when participants were presented with action verbs, their neural responses displayed significantly greater desynchronization of mu rhythms, which are known to be related to action or observation of it, in comparison to non-action verbs (Vukovic & Shtyrov, 2014). Additionally, this motor activation occurs in a somatotopic fashion, which means that processing a word associated with a specific body part activates a region related to that body part. For instance, the "hand" region of the motor cortex was triggered when participants were given an action word referring to hand activities such as "pick," while the "leg" area of the motor cortex was activated when they were provided with an action word referring to leg actions such as "kick" (Hauk et al., 2004).

Another line of evidence for the embodiment theory of cognition comes from the findings that activation of the motor areas conversely influences the process of language comprehension (Bergen et al, 2010; Chambers et al., 2004; Glenberg & Kaschak, 2002; Glenberg et al, 2008; Pulvermüller et al., 2005). The most direct evidence of this influence is obtained from the Transcranial Magnetic Stimulation study (TMS) (Pulvermüller et al., 2005), a method stimulating parts of brains by magnetic fields in order to improve or inhibit the activities of brain for experimental or clinical purpose. In this study, participants were asked to perform a lexical decision task that consisted of items with action verbs related to leg or arm actions. TMS stimulation was applied to the motor cortex associated with the arm or leg immediately after the word was presented. The results showed that participants responded more quickly to arm-related words when TMS was applied to the arm area, and vice versa when the leg area was stimulated. Although more indirect, this effect can also be supported by interference or facilitation effects between body-effectors. A number of behavioral studies have shown that comprehending words, phrases, or sentences that depict action (bodily movements) facilitates or interferes within the same-effectors or between different-effectors (Bergen et al, 2010; Chambers et al., 2004; Glenberg & Kaschak, 2002; Glenberg et al, 2008). For example, when participants listened to a sentence describing the leg action ("kick"), their response using legs became faster compared to when they listen to the one implying the hand action ("grasp"). It is assumed that comprehension process activates a certain body area of brain (here, leg area) and this activation helps facilitate subsequent movement of the same area by preparing it in advance.

The aforementioned findings provide strong support for the embodiment theory of cognition, as they align with the predictions of this theory. However, the disembodied theory disputes these results, arguing that they can also be explained from an amodal perspective. It argues that motor activation is considered to be a by-product of spreading activation from the amodal system to the sensory-motor areas (Mahon & Caramazza, 2008). When participants process action verbs, the amodal system is first activated and then motor system is activated through the mediation of the previous activation. This interface between amodal system and sensory motor system is similar to the concept of language faculty by Hauser et al. (2002) and in fact, it is a viewpoint held by most amodal perspectives, except for the extreme theory completely excluding sensory-motor information from language processing, which is rare.

In order for embodiment theory to overcome this criticism, it is necessary to provide evidence that activation of sensory-motor areas is not epiphenomenal spreading but an essential element of language understanding. One possible way of providing such evidence is through lesion studies that examine patients with brain damage. A lesion study is a method of investigating the causal role of a specific brain area by examining whether the function that is supposed to belong to the damaged area is impaired (Vaidya et al., 2019). For example, if conceptual processing of sounds from everyday items was impaired by damage to left auditory association cortex (Trumpp et al., 2013), it is inferred that the auditory area of sensory-motor cortex was crucial for sound-related objects. A numerous brain lesion studies demonstrated that lesions within each part of sensorimotor areas of brain resulted in poor comprehension of manipulable tools (Warrington & McCarthy, 1987), action-related verbs (Bak et al., 2011; Cotelli et al., 2006), visual information (Hart & Gordon, 1992) and acoustic concepts (Trumpp, Kliese et al., 2013). These findings suggest that the sensorimotor area is not epiphenomenally activated but plays a causal role in language processing.

While lesion studies provide valuable insights into the causal role of specific brain regions, they also have several limitations that restrict their widespread use. One of the main limitations is that the findings typically result from a small number of patients with specific lesions, due to the restrictive nature of the method. Since the structure of the brain does not vary widely among individuals, it is not that the findings from the lesion studies cannot be generalized, but it is still not sufficient to accumulate quantitative research that can be applied to more general situations. In addition, most studies have been conducted on patients with large-area brain injuries, which can make it difficult to pinpoint specific areas of the brain that are responsible for a particular function. Studies on patients with small lesions would be more effective in testing specific predictions about functional localization (Kiefer & Pulvermüller, 2012). However, since the number of patients with brain lesion itself is limited, it will be more difficult to find patients with more localized lesion in the specific area. Furthermore, the patients with larger lesion often mean a possibility that the function of an unwanted area is also impaired. Although other cognitive tests are also conducted to assess whether the loss of a certain function is due to overall cognitive decline or not, the absence of unmeasured effects still cannot be guaranteed and brain damaged patients tend to have inferior performance than non-patients in most tests.

In order to refute the claims of the disembodied view and offset the limitations of lesion studies, alternative experimental strategies are needed One potential approach is to increase the temporal resolution in order to examine differences in activation from early to late processing. If activation spreads from amodal system to sensory-motor system, it is assumed that the activation of amodal system occurs in earlier time windows than that of sensory-motor system. Therefore, if distinct neural activities are evoked in two different time windows and, moreover, if the involvement of the sensory-motor area is observed in later processing, this may prove the claim of the disembodied view. On the other hand, if the opposite trend is true, that is, if the activation of the sensory-motor area is obtained in the earlier time window, it would be supporting evidence for the embodiment theory.

In this respect, high temporal resolution methods such as EEG and MEG are more suitable than fMRI due to their ability to measure various temporal ranges with millisecond resolution. In fact, studies using EEG or MEG have provided support for the embodiment theory by demonstrating that the motor cortex is activated in early time windows during language processing (Cervetto et al., 2021; Dalla Volta et al, 2014; Harpaintner et al., 2020; Hauk & Pulvermüller, 2004;

Klepp et al., 2015; Vukovic & Shtyrov, 2014; Zhang et al., 2020). For example, in Harpaintner et al. (2020), from early time windows, EEG responses were distinguished between vision-related words and motor-related words as predicted in embodiment theory. The most commonly cited ERP components related to language processing are N400 and P600, which peak in ERP waveforms around 400ms and 600ms each, which suggests that semantic and syntactic language processing is integrated approximately around this time period. Also, other studies reported that semantic information is processed already within 250 msec after stimulus is presented (Kutas & Federmeier, 2011). Therefore, by investigating processing within 250 milliseconds before and after post-stimulus onset through EEG, the spreading activation problem can be partially overcome.

Another experimental strategy to rule out the assumption of spreading activation is to utilize an experimental paradigm that inhibits the propagation of activation. When a stimulus is perceived, various functional regions that cooperate to process that perception are activated, and they will subsequently spread to other related areas in a chain. However, by utilizing techniques such as masked priming, it is possible to elicit more localized activations and investigate processing in a more controlled manner. In a masked priming paradigm, a prime word supposed to affect a subsequent target word is exposed for only a short period of time, followed by a mask, which makes the prime unrecognizable to participants. Despite being unable to consciously perceive the meaning or presence of the prime word, it still has an influence on the processing of the target word in a subliminal, unconscious, and automatic way that is restricted to a lexical level (Forster et al., 2003). Thus, using masked priming can allow for the examination of lexical processing without the confound of epiphenomenal activation.

As described so far, the embodiment cognition account has been supported by accruing evidence of a range of methodologies from behavioral study to neuroimage study, which has shown activation of the sensory-motor system during language processing. While it remains a topic of debate whether this mechanism is essential for language comprehension or simply a byproduct of amodal computation, there are experimental methodologies and evidence that refute the disembodied view of language comprehension. Therefore, it can be reasonably inferred that language comprehension is at least partially rooted in embodied mechanisms and concurrently involves sensory-motor areas in conjunction with other systems.

2.1.2. Remaining issues on the embodiment perspective

While the embodiment cognition hypothesis has been supported by a growing body of evidence, there are still many areas that need to be researched and systematized in order to establish the theory on a solid foundation. First, since most of the previous studies have focused on the activation of motor areas (Borghi, 2020), it is necessary to expand the scope of the research to include a more diverse range of sensory modalities and features in order to fully understand the phenomenon of embodiment. As the embodiment theory originally emerged from the concept of mind-body-world interaction, the previous research has tended to focus on actions performed with the body. Many studies concentrate on how much motor area is activated by action words compared to non-action words under the dichotomy of “action *vs.* non-action,” subdividing only action words according to several body-effectors such as leg, arm, or mouth. However, this narrow focus on motor areas and action verbs is an overly restrictive approach, given the wide range of human perception. The emphasis on action is often so great that some theories even attempt to account for the concepts that are difficult to be explained under the term of action (e.g. abstract concepts) through the concept of action-motor activation assuming a metaphorical extension of action (Dreye & Pulvermüller, 2018; Glenberg & Kaschak, 2002; Glenberg et al., 2008; Glenberg et al., 2008; Moseley et al., 2012) or eventually have admitted that such concepts are not because they are not activated in the motor area (Mahon & Caramazza, 2008). However, human perception and interaction with the world are not able to be described only through motor-action. The sensory-motor areas encompass all the perceptions than action such as the input from human sensory organs and the information humans obtain while interacting with the world. Therefore, it is necessary to conduct research including various sensory modalities and abstract concepts that are difficult to elucidate under such a simple framework of action *vs.* non-action.

As embodiment cognition theory gains popularity, an increasing number of studies are exploring various features other than motor-action relationship. As a result, the range of concepts that can be explained under the term embodiment has been broadening. Some of them are: spatial terms such as upward or downward (Ahlberg et al., 2017), emotion (Dreyer et al, 2015; Dreyer & Pulvermüller, 2018; Pavlenko, 2012; Vigliocco et al., 2014; Zhang et al, 2020), sensory modalities (Harpaintner, Sim et al., 2020; Harpaintner et al., 2020; Hoenig et al., 2008; Kiefer et al., 2008; Kiefer et al., 2012; Kuhnke et al., 2020; Popp et al., 2016; Trumpp et al., 2013; Trumpp et al., 2014) and subclasses of abstract concepts (Borghi & Zarcone, 2016; Dalla Volta et al., 2014; Dreyer & Pulvermüller, 2018; Harpaintner, Sim et al., 2020; Harpaintner et al., 2020; Mazzuca et al., 2018; Muraki et al., 2020; Schalle et al., 2013). Despite this diversity of research domains, most areas are still not paid attention to by the field of embodiment cognition compared to motor-action. Therefore, there is a need to investigate a variety of modalities and features that have not yet been thoroughly explored. In this regard, this dissertation will focus on sensory modalities and abstract concepts that have been relatively under-researched.

As the scope of the embodiment cognition theory continues to expand, it is also important to broaden the study to include other languages and types of language users. Although the research on embodiment is conducted across various languages, the majority of studies have been conducted in European languages such as English, German, and Italian. Studies on Asian languages are substantially limited, and there are virtually no studies on Korean. Given that languages are embodied while interacting with the world, it is likely that differences in the environment surrounding languages will have a great impact on the embodiment of the language. For example, in the task of classifying action verbs by body-parts, participants associated different body-parts with each verb depending on the language they used to classify the words (Vasanta et al., 2011). Therefore, examining various languages and differences between them is critical to gaining insights into the embodiment cognition theory, especially when paying attention to the languages that have been neglected.

Another aspect worth of attention is embodiment processing of a different type of speakers—second language learners. Examining the embodiment of L2 processing not only provides additional information about embodiment theory, but also about L2 processing itself. Despite some previous studies on the embodiment of L2 processing, there is still limited understanding of this topic due to the paucity of accumulated evidence. This is not simply because of the small number of studies, but because of various individual variables in L2 that affect language processing such as in what environment L2 learners encountered the language, how proficient they are in L2 and when they started learning L2. The responses elicited by one group of L2 learners may not appear in the other group, and their responses may differ depending on the type of sentence employed in the task. Therefore, in order to investigate the embodiment of L2 processing, it is necessary to amass further evidence by taking various factors into account. This dissertation aims to contribute to the accumulation of such evidence by focusing on L2 learners who learn and use L2 in foreign contexts.

To sum up, the embodiment theory has accrued much evidence and has been able to refute the criticism of the theory in various ways, but many aspects of the theory still remain unsolved. First of all, it is necessary to explore various concepts that have been disregarded and underexplained by the theory which has focused on the motor area activation by action verbs. For example, abstract concepts are difficult to be explained by the involvement of the motor area, and investigations should be made for alternative explanations for such concepts. Lastly, the scope of research should be expanded to other languages and L2 language learners. Provided that the environment encompassing languages has a great influence on the embodiment of languages, focusing on the relatively neglected Asian languages and L2 learners will provide new insights into the theory.

In summary, the embodiment theory has amassed a significant amount of supporting evidence and has successfully addressed criticisms of the theory in various ways. However, there are still many unresolved aspects of the theory. One area that requires further exploration is the examination of concepts that have been overlooked or under-explained due to the theory's focus on motor area activation by action verbs. For instance, abstract concepts present a challenge to be explained by the involvement of the motor area, and alternative explanations should be investigated. Additionally, the scope of research should be broadened to include other languages and L2 learners. Given that the linguistic environment has a significant impact on the embodiment of language, focusing on understudied Asian languages and L2 processing may provide new and insightful perspectives on the theory.

2.1.3. The embodiment of abstract concepts

In order to fully comprehend the language representation within embodiment theory, it is necessary to thoroughly and systematically examine abstract concepts. Previous research has provided substantial empirical evidence for the embodied cognition approach to language, but this evidence is primarily limited to concrete concepts associated with motor-action. As such, using this framework alone may not provide sufficient evidence to explain other facets of language, such as abstract concepts.

Abstractness is defined as a concept or thought that lacks a tangible or physical presence and is derived from pure logic and unrelated to actual experiences or occurrences (Buccino et al., 2019). Intuitively, it is relatively easier to understand how concrete concepts, such as objects (e.g. hammer) or actions (e.g. kick), are processed through sensory-motor information. For example, when we think of a hammer, we can easily conjure up its shape, feeling, and arm movements, and this sensory-motor information may not vary greatly among individuals. If you describe what HAMMER means to a child, it might be better to show a picture of a hammer than to explain its definition in words. In contrast, it may not be straightforward to conceive how abstract concepts such as DEMOCRACY, HONEST, or POTASSIUM are grounded in our body or sensory-motor areas. Abstract concepts (although some concepts may not be regarded abstract such as POTASSIUM) do not seem to have their correspondent referent to the meaning. If you contemplate DEMOCRACY, no typical sensory-motor information may come to your mind, and if any, it might not be similar to that of the person next to you. Again, if you try to explain the meaning of DEMOCRACY to a child, showing an image of it, even if you can, may not be a smart way, and thus you need to somehow convey its definition with unsophisticated language.

The empirical evidence regarding the embodiment of abstract concepts in fact presents a significant challenge to the embodiment cognition theory. Some studies have failed to demonstrate the involvement of sensory-motor areas in the processing of abstract concepts (Dalla Volta et al, 2014; Desai et al., 2011; Wang et al., 2010), while others have reported that distinct areas of the brain are activated between when processing concrete words and abstracts words (Binder et al., 2005; Fiebach & Friederici, 2004; Holcomb et al., 1999; Kounious & Holcomb, 1994). For instance, Binder et al. (2005) showed that abstract concepts were left-lateralized similar to non-words, processed through ‘amodal’ linguistic modality, while concrete concepts activated both left and right hemispheres of brain which involves the imagery system (or sensory-motor areas in the terminology of the embodiment perspective). According to these findings, the scope of language that the embodiment perspective can account for is drastically limited to concrete concepts, placing the theory in an insecure state where it must remain silent about the concepts that make up a large part of language (Borghi, 2020). Then, as Mahon and Caramazza (2008) pointed out, why would we need such a specialized theory that can only give an explanation for concrete concepts such as objects and actions? Thus, providing evidence for the embodiment of abstract language is pivotal in the embodiment cognition research.

The embodiment of abstract concepts in language has been a contentious issue and various theories have been proposed to address this challenge. These theories can be broadly categorized into three groups. he first group posits that abstract and concrete concepts should be treated differently, often relying on amodal theory to explain abstract concepts. The second does not regard abstract concepts as entirely distinguished from concrete concepts, but it assumes that there are inherent characteristics of abstract concepts. This group of research attempts to find and employ the several characteristics only applied to abstract concepts in consideration of differences between abstract and concrete concepts, and claims that they are embodied through other brain areas than sensory-motor areas. The third group presumes that there is no qualitative difference in the embodiment of concrete and abstract concepts but only difference is the semantic complexity of concepts or the intensity of the activation. It adopts the same features and analysis method used for concrete concepts to investigate abstract concepts but considering the possibility of gradeability across the concepts.

The first category of research on abstract concepts posits that they are dependent on amodal, symbolic, or verbal representation under a representational dualism between abstract and concrete concepts. This dualism tradition can be traced back to the Dual Code Theory (Paivio, 1986), according to which cognitive processes involve two distinct systems: a verbal system in which a "logogen" is activated, and a non-verbal, imagery-based system in which an "imagen" is involved (Buccino et al., 2019). In the Dual Code Theory, it is argued that concrete concepts are represented through both imagens and logogens, while abstract concepts are involved only in the verbal system. This claim seems to be supported by the concreteness effect, which indicates that retrieving or recognizing concrete concepts is better and faster than abstract ones (James, 1975; Marschark & Paivio, 1977). It is assumed that concrete concepts are processed faster because they additionally activate imagery-based systems. In terms of neural activities, the Dual Code Theory predicted that abstract concepts rely on the verbal system and would activate the language-specific regions in the left hemisphere, whereas concrete concepts would activate both hemispheres including the visual imagery areas in the right hemisphere. However, the neurophysiological evidence for this is inconsistent. Some lesion (Funnell et al., 2001; Villardita et al., 1988) and neuroimage studies (Binder et al., 2005; Holcomb et al., 1999; Kounious & Holcomb, 1994) found a larger right hemisphere involvement for concrete words and a greater left hemisphere activation in language regions. Others failed to identify a distinct function of the right hemisphere in concrete word processing and instead observed the involvement of the sensory-motor areas for both abstract and concrete concepts or even greater right hemisphere activation for abstract concepts (Fiebach & Friederici, 2004; Grossman et al., 2002; Kiehl et al., 1999; Noppeney & Price, 2004; Pexman et al., 2007).

Despite ongoing debate over the degree to which the two hemispheres specialize in abstract or concrete language, the Dual Code Theory has had significant influence on research in embodied cognition theory during past decades

(Kiefer & Harpaintner, 2020). Some studies have taken a hybrid approach, considering abstract concepts to be inexplicable from the perspective of embodiment cognition and leaving their explanation to the amodal system (Andrews et al., 2014; Dove, 2009; Dove, 2011; Louwerse, 2008; Louwerse, 2018; Mahon & Caramazza, 2008). However, the interpretation of abstract concepts influenced by the Dual Code Theory is incompatible with the embodiment cognition approach. An argument that only concrete concepts are embodied is unreasonable and insufficient for the sake of theory. If abstract concepts that constitute about 70% of language (Lupyan & Winter, 2018) cannot be explained, the embodied cognition hypothesis will be of limited use as a theory. A more concise and plausible explanation would be, as proposed by disembodiment theory, that language is fundamentally processed by the amodal system but processing concrete concepts activates the sensory-motor area only secondarily just because of previous experiences with the physical entity they refer to and this activation at times facilitates the comprehension of concrete concepts as shown in the concreteness effect.

Indeed, this is how the Contextual Availability Model (Schwanenflugel et al., 1988) seeks to explain the concreteness effect. This model postulates the differences between concrete and abstract concepts stem from different degrees of accessibility to contexts stored in a single verbal system (Fiebach & Friederici, 2004). It assumes that understanding is primarily achieved by adding contextual information to the subject to be understood (van Hell & De Groot, 1998). This added contextual information comes from prior knowledge or the environment in which language is used. In contrast to concrete concepts which are closely linked to a physical referent, abstract concepts lack this information, and as a result, processing them is relatively more difficult. This hypothesis is supported by the findings that the concreteness effect disappeared when provided appropriate context to disambiguate the meaning of the concepts (Hoffman et al., 2010; Schwanenflugel et al., 1988; Schwanenflugel & Stowe, 1989; van Hell & De Groot, 1998) and the concreteness effect was even reversed, abstract concepts processed faster when they are more emotionally loaded than concrete ones (Kousta et al., 2011).

The Contextual Availability Model can be seen as closer to an amodal approach rather than an embodiment approach in that it proposes that all concepts are processed through a single verbal system. However, from the perspective of embodiment cognition it is worth noting the problems of the Dual Code Theory raised by this hypothesis. That is, the embodiment of language processing is not likely to be understood within the framework of the simple dualism of 'abstract *vs*. concrete.' The distinction between whether a particular concept is abstract or concrete is not clear-cut but rather gradual, variable and provisional. In fact, it is difficult to distinguish the concepts of various languages in a dichotomous frame of concreteness or imageability, the best-known measurements employed in psychology to categorize concrete and abstract concepts. When words are measured with such a criterion, there are always a number of concepts at the boundary between abstract and concrete. Moreover, even the same concept may be judged as abstract or concrete depending on the situation and context.

Despite these limitations, the dichotomy has had an implicit effect on embodiment cognition research. Under the influence of the dualism, research tends to be conducted in a way that alienated abstract concepts, whether intended or not. Numerous studies (Birba et al., 2020; Cervetto et al., 2021; Dalla Volta et al., 2018;

Dalla Volta et al., 2014; De Grauwe et al., 2018; Klepp et al., 2015) have used abstract concepts as a control group to prove the embodiment of concrete concepts, especially, action verbs. For example, in order to demonstrate that the sensory-motor area is involved during language processing, these studies compared the activation by action verbs with non-action verbs, mainly abstract concepts. It is an irony that this approach, while showing the embodiment of action verbs, indirectly supports the disembodiment of abstract concepts. Although they do not all claim that abstract concepts are processed in an amodal system their lack of explanation for abstract concepts ultimately excludes them from embodiment cognition theory. Therefore, in order to demonstrate language processing from the perspective of the embodiment cognition approach, it is necessary to provide a mechanism to expound the embodiment of abstract concepts away from the representational dualism influenced by the Dual Code Theory.

The second category of research, which does not accept the dichotomy between abstract and concrete concepts, proposes that abstract concepts activate some of the same areas as concrete concepts, but that the systems for them are not entirely overlapping, as abstract concepts have unique characteristics. For example, Barsalou and Wiemer-Hastings (2005) proposed that abstract concepts such as *truth*, *freedom*, and *invention* are embodied through introspective (mental, social, and affective) states. This group of research contends that previous neurophysiological studies have failed to identify an embodiment of abstract concepts because they do not take these unique characteristics into account. Therefore, this group of research aims to explore the embodiment of abstract concepts by identifying and examining these characteristics.

The Affective Embodiment Account is one of this type of research, which highlights the role of affective information in acquiring and processing abstract concepts (Kousta et al., 2011). It assumes that emotional words, which appear early in language development as emotional development often precedes language development in children (Ponari et al., 2018), play a role as "stepping-stones" in abstract semantic representations. Kousta et al. (2011) found a reverse concreteness effect, where abstract words are processed faster than concrete words, in contrast to previous findings. This reverse effect can be attributed to the greater association of abstract concepts with emotion, as the advantage for abstract over concrete concepts was lost when the emotional valences were balanced between them. These findings suggest that affective associations play an important role in the processing of abstract concepts. This conclusion is supported by an fMRI study (Vigliocco et al., 2014), which revealed greater activation of the rostral anterior cingulate cortex, an area known to be related to emotion processing, for abstract concepts compared to concrete concepts.

These findings demonstrate that emotion can be one of the important factors that distinguish abstract concepts from concrete concepts. However, they also showed that there were many abstract words that cannot be explained by affective association alone. As shown in Kousta et al. (2011), some of abstract words have low emotional valence. Altarriba et al. (1999) even differentiated the linguistic concepts using the trichotomy of concrete, abstract, and emotional. This indicates that emotion is one of crucial features of abstract words, but not all of them have a strong connection with it. Moreover, emotion is not encoded in specific neural circuits of the brain in one-to-one fashion between a certain emotion and a certain region (Buccino et al., 2019). It is argued that during processing emotions, as multiple systems are activated, emotion is also embodied just as language comprehension is. For this reason, it seems insufficient to explain the embodiment of abstract concepts only with the Affective Embodiment Account.

Another theory, Word As Social Tools (WAT), identifies several characteristics other than emotion as unique features of abstract concepts (Borghi et al., 2019). One of the most emphasized aspects of this theory is the heterogeneity of abstract concepts. Even though all abstract concepts appear to belong to the same abstract category against concrete concepts in terms of concreteness or imageability, simple reflection will lead to the conclusion that no abstract concepts are in fact the same. Just as concrete words *kick* and *punch* are classified into two different subgroups according to its body effector, *leg* and *hand*, abstracts words *think* and *angry* can be regarded as a word expressing *mental state* and an *emotion* each. Moreover, just as words with different body effectors are involved in distinct areas in the motor cortex, abstract concepts that belong to the different subgroups might activate different modality-specific areas in brain.

Previous studies have failed to acknowledge the diversity of abstract concepts and have instead used a singular measurement, such as concreteness or imageability, to classify them as "abstract," treating them as a homogenous group (Binder et al., 2005; Birba et al., 2020; Cervetto et al., 2021; Dalla Volta et al., 2018; Dalla Volta et al., 2014; De Grauwe et al., 2018; Fiebach & Friederici, 2004; Holcomb et al., 1999; Klepp et al., 2015; Kounious & Holcomb, 1994; Wang et al., 2010). This approach could be potentially leading to mislabeling, the concepts with different attributes as “abstract” or mis-grouping the concepts that consist of various features into “abstract.” For example, Dalla Volta et al. (2014) categorized 25 words as abstract concepts based on low concreteness and compared them with a group of words related to hand, foot, and mouth. From the findings that abstract words activated regions outside the motor system, it was concluded that abstract concepts could not be described with motor embodiments. Nevertheless, this conclusion has several limitations. First, the words classified as abstract concepts had also low hand-, foot-, and mouth-relatedness ratings. It is not unexpected that they failed to activate the regions pertaining to hand, foot, and mouth, given their low each body-effector-relatedness. Concrete words with low relatedness may not activate hand-, foot-, and mouth-related regions either. Furthermore, these words could be explained by an additional body-effector, head. Their head-relatedness was relatively high, 3.19 out of 5 on average (Lynott et al., 2020). If their body-effectors were categorized differently, the findings would have required a different interpretation. Finally, it is important to note that the words classified as abstract concepts in the study by Dalla Volta et al. (2014) were composed of 10 interoceptive words, 7 auditory words, and 8 visual words when evaluated by different sensorimotor criteria (Lynott et al., 2020). This indicates that while they were classified as a single group of abstract words, they in fact possessed slightly different characteristics. This can be problematic as neuronal activities are electrical signals that can cancel each other out when there are differences in timing, orientation, and location (Luck, 2014). For example, if interoceptive, auditory, and visual words are electrical signals originating from different regions, they may accidentally cancel each other out and create a null-effect. However, if they were all classified as abstract words and their neural responses were averaged, then a different conclusion could have been obtained from the findings. As in the example above, failing to account for the variability of abstract concepts might lead to mislabeling or misclassification errors, resulting in no effect and drawing inappropriate conclusions. WAT emphasizes the heterogeneity of abstract concepts in this respect, and moreover, assumes that since compared to concrete concepts, abstract concepts are more complex, more detached from physical experience and more variable, both across and within individual, in different contexts, rather more sophisticated classification is crucial (Borghi et al., 2019).

According to WAT, other than emotion, linguistic and social experience, metacognition, interoceptive signal, and mouth effectors characterize abstract concepts (Borghi et al., 2019). Although those characteristics also play a role in processing concrete concepts, it is assumed that abstract concepts are more engaged in them and the more abstract the words are, the more they are grounded in those features Of particular interest in WAT is the embodiment through linguistic and social experiences, as implied by the theory's name. This idea is inspired by the observation that abstract concepts are in many cases learned in social situation through linguistic information. The meaning of concrete concepts is easy to be acquired through interaction with real objects in the world, whereas abstract concepts are difficult to be obtained by direct interaction due to its absence of physical referent. Instead, they are experienced indirectly through verbal information or by the conventions of their usage within a specific society. In fact, the learning of abstract concepts is greatly restricted by the amount of social ability and linguistic experience in the developmental stage (Borghi et al., 2019). For example, it has been found that the acquisition of abstract concepts occurs relatively late (Ponari et al., 2018) and only after basic vocabulary has been acquired (Bergelson & Swingley, 2013) and social skills such as gaze following and joint action have developed (Brooks & Meltzoff, 2005; Buresh & Woodward, 2007).

The embodiment of abstract concepts through social experience is more easily understood in terms of situated simulations in the brain's modal systems (Barsalou et al., 2008). By reactivating the social situation or context in which concepts are acquired, they can be embodied and comprehended. On the other hand, embodiment through linguistic experience or linguistic mediation is difficult to be distinguished from the viewpoint of amodal theory. For example, according to distributional view (Burgess & Lund, 1997; Landauer & Dumais, 1997), which is contrasting to embodiment approach, word meaning is formed from a network of associated words, or statistics of how words are distributed—with which words a word frequently co-occur. That words are learned through other words seems similar to WAT's claim that abstract concepts are acquired through linguistic experience. However, WAT argue that it is differentiated from the distributional view in that it emphasizes that re-enactment of situated experience is required to comprehend the meaning (Borghi et al., 2019). In WAT, to use language as a medium is not to process the meaning of a concept through amodal symbols, but to reactivate the linguistic experience of when they learned it. Abstract words often have more complex meanings and unspecified referents, so that it is not feasible to reactivate the experiences from the referent in which meanings are grounded. Therefore, the meaning of an abstract concept is processed by re-enacting the linguistic experience of rehearsing the meaning of the word with language, as when the word was first acquired through language. It is hypothesized that the presence of inner speech reflects this re-enacting process. Inner speech is used to retrieve the meaning of words by retelling ourselves or asking others the meaning, which will be revealed by the activation of mouth region. As the processing of action verbs activates each body part referred to, abstract concepts are predicted to activate the mouth. This is supported by the findings that there was a correlation between abstract concepts and mouth activation in the word categorizing task where participants related abstract and concrete words to hand and mouth effectors (Ghio et al., 2013) and articulatory suppression led to slower responses for abstract words compared to concrete words (Fini et al, 2021).

It is worth noting that WAT highlights the diversity of abstract concepts and attempts to comprehend their embodiment through various characteristics related to them. It should be, however, acknowledged that this focus on heterogeneity is also a weakness of the theory. When a multitude of factors are proposed to explain everything, systematic explanation becomes more challenging and the explanatory power of the theory may be weakened, as new factors can be introduced as necessary. In order to explain abstract concepts, WAT introduced features such as emotion, linguistic and social experience, metacognition, interoceptive signal, and mouth effectors, and to evaluate words based on those features, 16 or 19 measurements were used (Villani et al., 2019; Villani, D'Ascenzo et al., 2021). Since each of these measures did not form an independent category, it was possible to group them to a smaller level through correlation analysis, (e.g., in Villani et al. (2019), the rating of *interoception* and *emotion* showed a high correlation of 0.8.) but it still seemed that too many factors were required to describe a single word. Furthermore, there were abstract concepts that were not accounted for even with these many features, such as numerical words which were regarded as an exception in WAT as they were not well classified as one of the significant features of abstract concepts considered by WAT (Myachykov & Fischer, 2019). If another concept is not justified in the existing framework, should another new feature be introduced? An overemphasis on heterogeneity greatly lowers the barriers to such adoption. It might be believed that if a concept is unexplained, it is due to the failure to consider the heterogeneity and thus the more fine-grained classification is required. In fact, some attempts are made to refine the classification even further, including not only institutional, theoretical, food and artefact, but also pure-institutional concepts and meta-institutional concepts (Villani, D'Ascenzo et al., 2021).

It is undeniable that the work of detailed categorization is valuable and contributes to expanding our understanding of language. In fact, to comprehend language processing, fine-grained analyses are essential. However, for this classification, it is crucial to first consider the measurability and neurophysiological aspects. For example, Villani et al. (2019) and Villani, D'Ascenzo et al. (2021) adapted the measurements of *social valence* ("how much the word evokes social circumstance") and *social metacognition* ("how much you think you have or needed others to understand the meaning of each") to measure the degree of *social experience* of a certain concept. The question arises as to how it can be confirmed whether the feature of *social experience* is really assessed by those two concepts, *social valence* and *social metacognition*. Moreover, the correlation between these two measures is 0.33, suggesting a positive correlation, but provided that the correlation between *abstractness* and *social valence* is 0.04 and the correlation with *social metacognition* is 0.5, those two may be measuring some distinct notions. If they measure rather different concepts, should each be considered a new feature? Or should *social experience* be viewed as a broader concept evaluated by more than two measurements?

The lack of measurability in classifying abstract concepts is often a result of basing categorization on the output or phenomenon of language. For instance, the characteristics of abstract concepts are searched for and classified through the observations such as “there are many abstract words related to institution” or “their acquisition occurs at a later age.” However, as for this type of classification, it is difficult to find a consensus among various theories and to find a neurological basis because it is not based on neurophysiological characteristics. In fact, the neurophysiological evidence for the abstract concepts in WAT seems to be unsatisfactory to explain the embodiment as it argues. Although it provides a great deal of neural evidence—left inferior frontal gyrus (mostly pars orbitalis, Broca area), left middle temporal gyrus and anterior superior temporal sulcus are involved in processing abstract concepts, which are known to be related to the activation of language processing or social processing— most of its evidence is consistent not only with WAT’s claim that the linguistic experience is significant but also with the amodal theory or hybrid approach, since they also predict the involvement of the so-called linguistic areas. So, WAT is in fact indistinguishable from the hybrid approach or amodal theory in terms of neurophysiological basis.

In addition, the brain regions related to social processing are not limited to a specific part of the brain, but the activations occur across various parts such as the medial prefrontal cortex, posterior cingulate cortex, amygdala, right fusiform gyrus, bilateral temporal pole, and right inferior frontal gyrus (Tso et al., 2018). Because social experience is also obtained through complex processing like language comprehension, it might be not possibly classified into a single category. Perhaps the characteristic of social experience itself also has heterogeneity. In this respect, when defining the characteristics of an abstract concept from the point of view of embodiment cognition, it is necessary to classify them, focusing on neuroanatomical substrates and neurophysiological dynamics of cognitive processes rather than based on the output or phenomena of language (Mkrtychian, et al., 2019).

The third group makes no distinction between concrete and abstract, arguing that they are all embodied in the same way. From this point of view, it is not necessary to distinguish between concrete and abstract concepts because a linguistic concept itself is abstract. In the classic British empiricist tradition, words are largely divided into proper names and general terms (Buccino et al., 2019). All general terms, whether referring to "chair" or "virtue", are not intended to refer to a particular object. What "chair" refers to is not always the same under any circumstance or to any person. Even “chair,” a concrete word, is something in which world experiences are abstracted. However, it is assumed that the concepts classified as abstract words differ from concrete ones in the degree of abstraction or the complexity of the underlying experience. Abstract concepts appear to be perceived differently from concrete one because they are more complex (Buccino et al., 2019). The complexity of abstract concepts can be viewed, in a broader sense, as similar insights, such as heterogeneity of abstract concepts in WAT or lack of contexts assumed by the Context Availability model. However, this view is distinguished from those in that it regards concrete and abstract as not fundamentally different and places emphasis on the continuum from concrete to abstract words. From this point of view, there is no need to reveal unique properties of abstract concepts. They are just more complex, and thus they can be delineated by employing a similar research method used for concrete concepts. This view has the advantage of being concise in the logic and methodology of the theory. Abstract concepts can be explained also in terms of the activation of sensory-motor areas without introducing new characteristics only for them, such as linguistic or social experience, discussed in WAT.

The complexity of abstract concepts is characterized by their modal-unspecific or multi-modal nature (Buccino et al., 2019). The meaning of an abstract concept is more likely to be grounded in multiple modalities or effectors rather than just one. For example, the concept of “democracy” can be grounded in different body-effectors and modality, depending on which experience is re-enacted. Hand-effector and haptic sensory can be activated in reference to the experience of voting or mouth-effector and auditory sensory to speech, or arm-effector and emotion-related areas to protest. In contrast, the semantics of concrete concepts are expected to be grounded in a relatively typical effector or modality. A concrete concept such as “walk” may activate only a foot-effector while “ice-cream” gustatory sensory. In fact, when evaluating words based on sensory modalities—how much each sense is used to understand the word, or perceptual strength—participants felt more difficulty in recognizing and evaluating some abstract words (Lynott et al., 2020). In addition, there was a weak but positive correlation between concreteness and exclusivity, which indicates how much a word can be expressed by one specific sense (Brysbaert et al., 2014; Lynott et al., 2020). This suggests that abstract words are more likely to be associated with multiple sensory modalities. However, it is important to note that concrete concepts can also be multi-modal, depending on various factors. For example, "ice cream" can be associated with the gustatory, olfactory, and visual modalities.

The meaning of abstract concepts is also more complex as it is dynamic. The modality in which an abstract concept is grounded can vary depending on the specific experience being reproduced, and that experience can differ among individuals and across different contexts. For example, the concept of "democracy" may evoke different experiences for those who live in a society where democracy is not guaranteed, those who have struggled to achieve it, and those who do not participate in acts such as voting. Additionally, the nuances of the concept may also vary depending on factors such as the speaker's age, education, or lifestyle, as well as the historical time or environment in which they are speaking (Buccino et al., 2019). For instances, Villani, D'Ascenzo et al. (2021) demonstrated that the degree of embodiment may be different for each individual’s experience, where legal experts showed more embodied responses on word categorization, using the emotional experience of institutional concepts. In this respect, abstract words are less fixed compared to concrete words. But it is necessary to keep in mind that the concrete concept may not be fixed as well, since the degree and target of embodiment may differ among individuals, just as feelings for *dogs* differ from person to person.

The current neurophysiological evidence has not yet provided substantial support for the third perspective on embodiment cognition. Studies conducted according to the Dual Code Theory have consistently demonstrated that concrete and abstract concepts are processed independently in the brain (Binder et al., 2005; Holcomb et al., 1999; Funnell et al., 2001; Kounious & Holcomb, 1994; Villardita et al., 1988). Even in WAT, which attempts to explain abstract concepts from the point of view of embodiment cognition, it is considered that although abstract and concrete concepts are not dichotomously classified, they are in general processed separately (Borghi et al., 2017; Borghi et al., 2019). However, as shown in the reverse concreteness effect (Hoffman et al., 2010; Kousta et al., 2011; Schwanenflugel et al., 1988; Schwanenflugel & Stowe, 1989; van Hell & De Groot, 1998), some findings contradict the distinct neural activation patterns between abstract and concrete concepts. Indeed, several neurophysiological studies have found no significant differences in brain activation patterns between abstract and concrete concepts (Fiebach & Friederici, 2004; Grossman et al., 2002; Kiehl et al., 1999; Noppeney & Price, 2004; Pexman et al., 2007). Moreover, as mentioned above, the findings from the previous studies that dichotomously divide abstract and concrete concepts, regarding abstract concepts as a homogeneous group, may not be reliable in terms of the embodiment cognition. The results were likely to be derived from the mislabeling or misgrouping of several different concepts. For more accurate results, detailed classification of abstract words using acceptable criteria is required.

In terms of classification, one advantage of the third perspective is that it does not need to introduce a new classification system only for abstract concepts as in WAT. Instead, it categorizes them based on the sensory-motor properties in the same way that the concrete concepts are classified. Perceptual strengths classified by such criteria were one of the most important explanatory mechanisms to explain the response times of lexical decisions tasks. (Lynott et al., 2020), which were often used to account for the concreteness effects. In Lynott et al., (2020), Minkowski 3 distance was the factor that best predicts the reaction times of lexical decision tasks, which is similar to a weighted sum of perceptual strengths across all modalities, indicating that the concreteness effect can be better explained when multi-modal characteristics were taken into account. Additionally, this categorization was supported by neurophysiological evidence, where both concrete and abstract concepts were embodied according to the sensory-motor characteristics regardless of their concreteness. In Harpaintner, Sim et al. (2020), when abstract concepts were divided into motor-specific and visual-specific words, the former elicited increased activation in motor regions, whereas the latter activated visual-related areas. This shows that abstract concepts also engage sensory-motor areas in the same way as concrete concepts when properly classified according to the sensory-motor properties.

Classification systems using the sensory-motor characteristics generally focus on several body-effectors and five common sensory modalities—vision, audio, touch, smell and taste. However, it has been suggested that this classification alone may not be sufficient to fully describe the embodiment of abstract concepts, and therefore it may be necessary to introduce an additional sensory modality—*interoception*. It is one of most important senses for human beings, but it is often forgotten in the sensory-motor classification. Interoceptive sensation refers to perception to the internal state of the body, which includes muscular and visceral (cardiovascular, respiratory, gastrointestinal) sensations, vasomotor activity, hunger, thirst, breathlessness, pain, pleasure, temperature, itch, and sensual touch (Craig, 2003; Craig, 2011). These sensations are known to play an important role in emotional processing and maintaining physiological homeostasis by signaling changes of individual physiological parameters (Connell et al., 2018; Critchley & Harrison, 2013). Since interoception is crucial to human beings and widely experienced although not always consciously attended to it, it should be regarded as an essential characteristic of sensory-motor perception. One might argue that the sensations associated with interoception are so broad that it must be distinguished from other senses and may not be classified into a single category. However, a wide range of sensations also exists in other modalities such as vision and touch. For example, touch is a sense that incorporates pressure, texture, movement, vibration, and tactile cold/heat (Connell et al., 2018). Also, various aspects of interoceptive sensation have a common neural mechanism and can be measured in a similar way to other modalities (Critchley & Harrison, 2013). Interoception is a particularly important sensory modality for studying the embodiment of abstract concepts, because it plays an important role in the processing of emotions (Connell et al., 2018; Critchley & Harrison, 2013). Indeed, Lynott et al., (2020), which measured the perceptual strengths of 40,000 words including interoception, found a negative correlation between interoceptive ratings and concreteness (See also Villani et al, 2021). This supports the idea that interoception should be considered as an additional modality in the classification system for embodiment cognition.

In summary, the question of how abstract concepts are embodied in language remains a topic of ongoing debate, with various theories proposed to address this challenge. The first group posits that abstract and concrete concepts are separately processed in the brain and that it is appropriate to explain abstract concepts from an amodal perspective. This theory has been widely influential and supported by experimental evidence, but it is at odds with the embodiment perspective. The second group does not rely on the dichotomy of abstract and concrete concepts, but acknowledges the heterogeneity of abstract concepts. Heterogeneity is obviously an important property of abstract concepts, but an overemphasis on it leads to overly subdivided and unstructured categorization on abstract concepts. The third group argues that there is no distinction between abstract and concrete concepts, and that abstract concepts can be understood through the activation of the sensory-motor domain. While this approach has a logical simplicity, it is deemed as an overly strong form of embodiment cognition and has limitations in explaining some of current neurophysiological data.

This dissertation will focus on the perspective of the third group of theories proposed to address the embodiment of abstract concepts in language, while acknowledging the claims of the second group as well. The third group's argument is considered to be the most straightforward as it suggests that abstract concepts can be described using the existing embodied cognition framework. Despite evidence suggesting that activation of sensory-motor areas alone cannot fully explain the embodiment of abstract concepts, including interoception in the sensory-motor system may refute such evidence. However, the question of how to account for the activation of language-specific brain regions during abstract word processing remains unresolved, which is key evidence supporting the role of linguistic and social experiences in the embodiment of abstract concepts. Therefore, as suggested by the second group, the involvement of areas outside the sensorimotor domain should not be completely disregarded.

2.1.4. Embodiment in L2 processing

The present study also raises the question of whether the language comprehension process of L2 learners is "embodied" in the same way as L1 learners. The learning environment for L2 adult learners, in terms of time spent using the language, the quality of input, and the context of language access, differs from that of L1 child learners. When children acquire their native language, they encounter the language in numerous contexts. For example, assuming an environment in which a child learns a word, “apple,” he or she learns the word by seeing, tasting, and feeling apple, that is, associating the context of all senses to the concept named “apple.” In contrast, in an adult speaker's L2 learning environment, learners have significantly less opportunities to engage in the sensory context surrounding words. For general cases, they learn words mediated through their own L1 translate equivalents. In this respect, the sensory-motor areas may not be involved in L2 processing, especially considering the mechanism of associative learning, which is regarded as a key principle underpinning brain structure from the embodiment theories (Matheson & Barsalou, 2018). According to associative learning theory, it is assumed that repetitively associating words and behavior or experience forms the link between them (Pulvermüller, 2005). This is based on the Hebbian theory, which states that the simultaneous activation of separate cells can generate a functional unity (Hebb, 2005). If two neurons are stimulated periodically at the same time, they strengthen their bond. As a result, when one of these two neurons is activated, the other one is activated as well. For instance, Pulvermüller (1996, 1999) suggested that words often used in physical activity contexts (e.g., action verbs) would be coded into both language and motor area. In this respect, the embodiment of language is constructed by human cognitive experience in associated language learning environment, and thus language may not be embodied in L2 learning environments due to less physical contact experience.

However, regardless of differences in environment and the number of contacts, L2 learners also confronts the instances in which language and sensory context are connected. In fact, several studies indicate that language can be embodied even with a limited amount of sensory-motor experience. For example, Beauprez et al. (2020) has found that even short observations can lead to activation in the sensory-motor domain when participants read sentences describing unusual actions. Furthermore, even assuming that L2 learners understand L2 entirely relying on L1 as a medium, their L1 is still rooted in the sensory-motor system, which may in turn have influence on L2 processing. Günther et al. (2020) showed that words can be grounded in sensorimotor area without direct experience but only with indirect experience mediated through language. In the experiments, participants' movements were facilitated while reading sentences including novel words implying the same movement even after they learned those words only by their definition. Based on these findings, it is also predicted that L2 processing is to some degree embodied as with L1.

Although much scanter compared to L1 studies, the question of whether L2 processing is embodied has been increasingly discussed in recent years. Research on the embodiment of L2 processing can be divided largely into two groups: the studies on a lately-acquired second language and on a newly-acquired language (novel word paradigm). Frist, evidence from lately-acquired second language has been obtained from several behavioral and neural studies. In Bergen et al. (2010), after L2 learners of English (mostly Japanese native speakers) saw a picture describing a particular action, they responded slowly when processing English verbs that shared the same effector with the picture (e.g., hand action verb for hand action). Buccino et al. (2017) and Vukovic (2013) also showed similar interference effect within the same body-effector among L2 participants. In particular, in Vukovic (2013), the interference effect was mediated by proficiency, where only highly proficient learners responded slowly to the same effector verb. Additionally, neural evidence for the interference effect has been found through EEG studies, showing greater desynchronization of mu rhythms related to observing action when L2 learners processed action verbs (Vukovic & Shtyrov, 2014). The involvement of sensory-motor area during L2 processing was also found in fMRI studies. In De Grauwe et al. (2014), German learners of Dutch performed a lexical decision task including Dutch motor verbs and non-motor verbs. When processing motor verbs, there was stronger motor area activation comparable to those from Dutch native speakers. Embodiment effects appeared not only in action verbs, but also in word stimuli representing spatial relationships (Ahlberg et al, 2018; Dudschig et al., 2014; Vukovic & Williams, 2014). Ahlberg et al. (2018) conducted a Stroop task for German native speakers and several groups of non-native speakers (English, Russian, Korean, and Turkish), where the participants were required to press either an upper button or a lower button depending on the presented font color (e.g. blue for upper button and red for lower button), which was compatible or incompatible with the presented spatial preposition words. For example, the upward response to *über* (above) or *auf* (on) is a compatible condition, while the upward response to *unter* (under / below) is an incompatible condition. The findings revealed that L2 participants responded faster under compatible conditions regardless of their first language, as L1 speakers did. However, L2 learners did not show a difference in compatibility effect between *über* and *auf*, which was different from L1 speakers. Since *auf* implies a smaller distance of movement on the vertical axis and non-spatial use is more common than *über*, smaller compatibility effect was predicted, and in fact, there was no compatibility effect in *auf* for L1 speakers. To summarize, the studies on lately-acquired second language have shown that L2 processing is also embodied. Some of them demonstrated that there is a difference in the degree of embodiment between L1 and L2, while others found a variation by L2 proficiency. Most of the studies were conducted on motor verbs, but some also utilized spatial prepositions as stimuli.

Various studies have explored L2 embodiment processing using novel word paradigm. This approach allows for easy control of various variables, such as learning histories, and addresses the question of whether even newly acquired words can be embodied. Many studies have sought to prove the embodiment of newly-acquired language by examining whether word learning accompanied by gesture induces better recall. When the words of artificial languages, Vimmi and Tessetisch, were learned for a short period of time (2 to 5 days), the words taught with a semantically congruent gesture were recalled and translated better than the incongruent condition (Garcia-Gamez & Macizo, 2019; Macedonia & Klimesch, 2014; Macedonia & Mueller, 2016; Macedonia et al., 2011; Mayer et al., 2015). This motor-language coupling effect was found to last up to 14 months (Macedonia & Klimesch, 2014). In addition, the recognition of words learned along with gestures activated sensorimotor areas of brain (Macedonia & Mueller, 2016; Macedonia et al., 2011; Mayer et al., 2015). This embodiment occurred not only when the participants actually performed the gesture, but also when it was observed or paired with an image representing it. For example, when native speakers of English learned Japanese verbs along with observation of iconic gestures, they showed faster word recognition and more accurate translation than those without, resulting in different EEG responses (late positive complex) related to language processing (Kelly, McDevitt & Esch, 2009). In summary, word learning combined with gesture or observation resulted in improved recall of novel words in the newly acquired language despite only a brief duration of contact and the involvement of the relevant sensory-motor region during word recognition.

Despite the advancements made in understanding the embodiment of L2 processing, several questions remain unresolved. One significant limitation is the limited number of studies that have been conducted in this area. Research on neural responses, specifically, is scarce, and when grouped by specific techniques such as EEG and fMRI, the number of studies in each group is even fewer (Kogan et al., 2020). This lack of research makes it difficult to fully understand the neurological basis of L2 embodiment. Moreover, many of these studies have been conducted on artificial or newly-acquired languages. A significant portion of the research analyzed in a recent meta-analysis of L2 embodiment (Kogan et al., 2020) deals with studies using the novel word learning paradigm. This paradigm clearly has the merit of allowing researchers to investigate how an embodiment can occur in a short amount of time, as well as the dynamic of neural substrates in the early learning process, by studying the immediate effect of learning and taking control of a variety of learning environments. However, the novel-word paradigm also poses certain limitations. The immediate effect of learning in this paradigm, as it is more likely to be a short-term effect, may not be congruent with the real-world language learning and recognition process. Since knowledge of words can have multiple layers, the representation of a word acquired in the short period of time may not be identical with that of a word stored in our mental lexicon. It is difficult to predict what changes will occur when a word is learned and that knowledge matures. In particular, the representation of L2 word, due to the interaction with the L1 translation equivalent, might be altered dynamically over time, which is difficult to capture by the newly-acquired language paradigm. Furthermore, the majority of these paradigms are based on learning accompanied by gestures, which may not fully capture the diverse characteristics and learning processes of actual L2 language learning environments. As a result, it can be argued that the gesture-based paradigm provides only a limited understanding of the actual L2 language learning environment.

The studies on lately-acquired L2 languages are even fewer which makes it difficult to take into account the various individual variables that impact L2 processing, including variations in learning environments. Although it is difficult to investigate all of the elements that impact L2 processing in one study, it is vital to gather data by performing research on each variable in each study. Additionally, there has been a lack of effort in defining the distinction between L1 and L2 in previous studies. Most studies have focused solely on whether L2 processing is embodied or not and many of them have been carried out without L1 control (Buccino et al., 2017; Ibáñez et al., 2010; Vukovic & Williams, 2014; Xue et al., 2015). However, as shown in the research comparing L1 and L2, there may be variances between the embodiments of L1 and L2. By exploring these differences, it is expected that a deeper understanding of embodiment and L2 processing can be gained, so comparison with L1 is essential.

Lastly, previous research on L2 embodiment has primarily focused on words describing concrete concepts, such as motor actions, similar to the majority of L1 literature. However, L2 embodiment may vary depending on the linguistic aspect under investigation. For example, research has shown that there is little to no sensory-motor involvement in L2 processing of emotion-related language (Pavlenko, 2012). Thus, in L2 embodiment cognition research, it is also necessary to extend the scope of the research to including abstract concepts. L2 research on abstract concepts, including emotional concepts, can enrich current discussions about abstract concepts. Linguistic and social experiences are thought to be crucial for abstract concept in WAT. Since there are many cases of learning L2 through language in the L2 learning environment, L2 is more likely to rely on linguistic and social experience for processing abstract concepts. Furthermore, in contrast to L1, linguistic experience may be essential in the processing of concrete concepts as well as abstract ones if L2 is learned mostly through the mediation of language. As the complexity of abstract words, according to WAT, is a major reason why language and social experiences are used to process them, the concepts that are difficult for L2 learners to acquire, independent of concreteness, may be embodied through verbal and social experiences.

In sum, previous studies have shown evidence of the involvement of sensory-motor areas in L2 processing. However, such evidence has numerous limitations. First, there are few research, and there are even less studies on neural responses. The majority of the research focuses on newly-learned language, with motor-action verbs serving as the primary stimuli. Furthermore, several studies did not make any comparisons with L1 controls or did not take into account distinct individual L2 processing characteristics. This study aims to conduct research by attempting to address these shortcomings of previous studies.

2.2. Related works in Second Language Acquisition

2.2.1. Emergentist models

Emergentist model is an umbrella term for the views that language learning arises from language use and experience, rather than from innate language systems (Gass et al., 2020; Gregg, 2003). This perspective encompasses usage-based model, frequency-based model and connectionist model (Gass et al., 2020). In contrast, generative linguistic theory assumes that language acquisition is based on mental representations of innate language systems or syntax/grammar (Ellis, 1998). From this point of view, language knowledge is not acquired from the process of learning experiences but is derived from Universal Grammar (O'Grady, 2008). This argument stems from the observation that a logical flaw in the language acquisition process, namely, that children obtain their linguistic achievement in spite of poor linguistic input (Ellis, 1998). Our linguistic knowledge often goes beyond the input of our linguistic stimuli, and people can produce sentences they have never heard of and judge a sentence to be ungrammatical even if they have never been taught it. On the other hand, Emergentist model argues that language learning is not originated from the innate language system, but general learning mechanisms, whether it is first, second or third language acquisition, which is the ability to extract structures and patterns from language stimuli (Gass et al., 2020). Human beings have a learning mechanism that obtains meaningful patterns from environmental inputs, including all kinds of sensory perception, and this capacity is also used to acquire languages (Ellis & Wulff, 2015). Although this model stresses the importance of input, the ability to extract probabilistic patterns for input is essential; hence, it is different from behaviorism that learning is accomplished merely through input and error correction (Ellis & Wulff, 2015). The observation that people can speak sentences they have never heard of can be explained by probabilistic interpretation and application. From this perspective, language acquisition is learning a pattern or construction between form and meaning/function (Gass et al., 2020). The constructions exist at all aspects of language from words to sentences, and the simplest form of constructions is a word, a connection between sound and meaning (Ellis & Wulff, 2015). The learning of construction is achieved through the learning of associations between form and meaning. The more reliable the association between a form and meaning, the easier it is to learn it. The reliability of an association is determined by various factors such as frequency, salience, redundancy, and context (Ellis & Wulff, 2015). For example, if an object appears frequently and repeatedly with a certain sound, the sound and object will be considered as a highly reliable association.

Emergentist models presume the same language learning mechanism for both L1 and L2; hence, L2 learning is equally achieved through probabilistic association of input, and this association is equally influenced by frequency, salience, redundancy, and context (Ellis, 2012). However, L2 learning differs from L1 learning due to the difference in the amount of input and the influence of the learners' native language. Since the input of L2 is qualitatively and quantitatively inferior to that of L1, it is insufficient for successful L2 learning (Ellis & Wulff, 2015). Poorer input may impede the formation of a complete association or result in probabilistic tuning that differs from L1. In probabilistic association, the influence of the native language occurs at all stages, and emergentist models consider this to be a major factor contributing to difficulties in L2 learning (Ellis, 2006; Ellis & Wulff, 2015). Although the influence of learners' native language is not as decisive a factor in learning as asserted by Contrastive Analysis, its influence, whether positive or negative, cannot be denied (Ellis & Wulff, 2015). One of the negative effects is that structures already established through native language learning interfere with the intake of L2 inputs (Ellis, 2006; Ellis, 2012). For example, L2 prefers structures that are salient, or namely structures that exist in their native language, due to learned attention, suppressing the intake of a redundant structures with similar meaning (Ellis, 2006; Ellis & Wulff, 2015).

Embodiment cognition approach regards associative learning as one of its principles, which posits that language is learned by the association of experiences with language and environment rather than an intrinsic amodal linguistic system. This learning mechanism is similar to that of Emergentist Model. However, the embodiment cognition approach differs from the Emergentist model in that it emphasizes the interplay between language and the environment. Although the Emergentist Model does not rule out the embodiment of languages, it is more concerned with interactions between and within languages, or how often and in which linguistic contexts a certain structure is used.

While the assumptions of the Emergentist model may not be fully applicable to the Embodiment Cognition theory, some of the model's observations regarding L2 learning may be relevant to the embodiment of L2 processing as both theories share the principle of associative learning. As suggested by the Emergentist Model, the relatively poor input from L2 learners' environment might cause the embodiment of L2 processing to differ from that of L1, making some aspects of L2 learning more difficult. As Emergentist Model considers crucial the effect of L1 on L2, especially interference caused by the established association of L1, in the embodiment of L2 processing, it is also necessary to pay attention to the influence of the learner's native language. In terms of embodiment cognition, the influence of already established structures is not necessarily negative in that both L1 and L2 mostly refers to the same referent in the world. A pre-existing construction between the learners' L1 and referents may facilitate forming a relationship between L2 and the world, compensating for L2 learners' lack of contacts with environment. However, since L1 and L2 do not have exactly the same referent due to cultural or linguistic differences, the experience of the native language may operate as interference. L2 learners may embody language differently than L1 speakers, and only some salient objects such as concrete concepts can be embodied.

2.2.2. Lexicon

This dissertation focuses on vocabulary processing among various aspects of language. As the application of Embodiment Cognition theory to L2 processing is relatively limited, this research investigates language processing at the lexical level as an initial stage study. This chapter provides an overview of how the mental lexicon has been studied within the L2 research tradition and how this relates to or is compatible with the Embodiment Cognition theory.

2.2.2.1. Changing views on lexicon

In early theories of language acquisition, language was conceptualized as a set of grammatical rules, and words were primarily seen as filling gaps in sentences created by syntactic transformations. Consequently, the mental lexicon was viewed as nothing more than a collection of lexical, phonological, and morphological information related to words (Dóczi, 2019). From this perspective, vocabulary capacity was not given sufficient attention when discussing the distinction between L1 and L2 acquisition. For example, the claim that late L2 learners would never attain the level of native speakers in the Critical Period Hypothesis was primarily based on qualitative differences in syntactic ability or pronunciation (Johnson & Newport, 1991; Long, 1990).

In recent years, however, there has been a shift in perspective regarding the importance of vocabulary knowledge in language acquisition (Webb, 2019a). Words are now recognized not only as filling grammatical structures, but also as a basis of grammar, as the distinction between vocabulary and grammar has become increasingly blurred (Gass et al., 2020). Usage-based approaches propose that the learning of syntax occurs in a similar manner to, or in conjunction with, the learning of words (Ellis & Wulff, 2015). From this perspective, constructions refer to the association between form and meaning, which encompasses not only individual words, but also phrases and sentence structures. Furthermore, learners tend to view errors in word usage as the most serious type of errors, with vocabulary errors being the most prevalent among learners. Anecdotally, errors in word usage are more serious than other errors because they can lead to semantically non-inferable results. As a result, along with the change in perspective on language, there has also been a shift in perspective on the mental lexicon and vocabulary learning.

2.2.2.2. Bilingual mental lexicon models

Mental lexicon refers to the cognitive system that constitutes the capacity for conscious and unconscious lexical activity (Jarema & Libben, 2007). It is unclear whether the mental lexicon exists in the brain as a separate module, but it serves as a metaphor to help us understand word storage, retrieval, and acquisition (Dóczi, 2019). A number of models have been proposed to explain the organization of the mental lexicon, including the hierarchical network model, semantic feature model, and spreading activation model. The *hierarchical network model* (Collins & Quillian, 1969) posits that words are interconnected in a hierarchical manner, with higher-level categories containing more general concepts and lower-level categories containing more specific concepts. For example, “animal” exists in the upper category, “bird” or “fish” in the middle category, and “penguin” or “salmon” in the lower category. *The semantic feature model* (Smith et al., 1974) suggests that concepts are stored in our memory as a set of properties. “Bird” is a set of “feathered”, “winged” and “can fly”, and “ostrich” is composed of “has feathers” and “long neck and legs”. Words that share more properties with each other are connected by a strong connection, otherwise the connection to each other is considered weak. *The spreading activation model* (Bock & Levelt, 1994; Collins & Loftus, 1975) presumes a lexicon as a network of associations, with concepts represented as nodes that are interconnected based on the strength of the associations between them. For example, “red” is connected with “orange, cherry, fire, rose” and is understood through those connections. Of these models, *the spreading activation model* is the most widely accepted because the results of the psycholinguistic studies, particularly the ones from the priming tasks, showed that words were related to each other, establishing a network of associations (Dóczi, 2019).

The bilingual mental lexicon model can be broadly divided into two categories. One argues that the L1 and L2 lexicons are completely separated, while the other claims that the two lexicons share concepts and interact with each other even if they are not completely integrated. The latter being currently accepted more widely, various models exist for how much two lexicons are integrated and how they interact (Dóczi, 2019). Mental lexicon models that assume the shared concepts of L1 and L2 lexicons are further divided into four types: *word association model, concept-mediation model, mixed model,* and *revised hierarchical model* (French & Jacquet, 2004) (Figure 2.1). *The word-association model* posits that L2 lexical items are directly linked to the translate equivalent of L1, but not concepts. Therefore, when L2 learners process a word, the concept is represented through the medium of their first language. This model is often considered to be suitable to explain the mental lexicon for less proficient learners. *The concept-mediation model* proposes that both L1 and L2 are connected to the concept, but not to each other. In view of this model, it is possible for language users to process words in each language without the intervention of a translation equivalent. It is assumed that higher level learners will shift to this type of mental lexicons as the connection between L1 and L2 gradually weakens. *The mixed model* is a combination of the above two models, and hypothesizes that both L1 and L2 are connected to the concept and are also connected to each other. This model predicts that increasing proficiency does not eliminate the link between L1 and L2 of the mental lexicon. *The revised hierarchical model* has the similar assumption with *the mixed model*, but differentiates L1 and L2 in connection strength. It is assumed that the connection between L2 and concepts and from L1 to L2 are relatively weak, compared to the connection between L1 and concepts and from L2 to L1.

The four basic models of bilingual lexical organization outlined previously are simplified versions of the bilingual mental lexicon model. Lexicons closer to actual data are more complex, so other models that reflect that complexity have been proposed: *the shared asymmetrical model* (Dong et al., 2005; Figure 2.2) and the *modified hierarchical model* (Pavlenko, 2009; Figure 2.3). They especially differ from the four models in that they assume three categories of concepts: common elements, L1-specific elements, and L2-specific elements. This distinction is considered to be a more accurate representation of the data, as L1 translation equivalents do not typically have a one-to-one correspondence with L2 words. For example, some words, particularly concrete words, may be represented in shared concepts, while others may have partial overlap or be represented in only one language-specific concept. Both models assume that the L2 lexicon is connected to all three elements, but they disagree as to which concept is most closely connected. The *shared asymmetrical model* postulates that the lexicon of L2 is more closely related to L2 specific concepts through response time analysis (Dong et al., 2005). In contrast, the *modified hierarchical model* posits that L2 lexicons are primarily associated with L1-specific or shared concepts and extended to L2-specific components by restructuring when L1-specific or shared concepts are not available (Pavlenko, 2009).

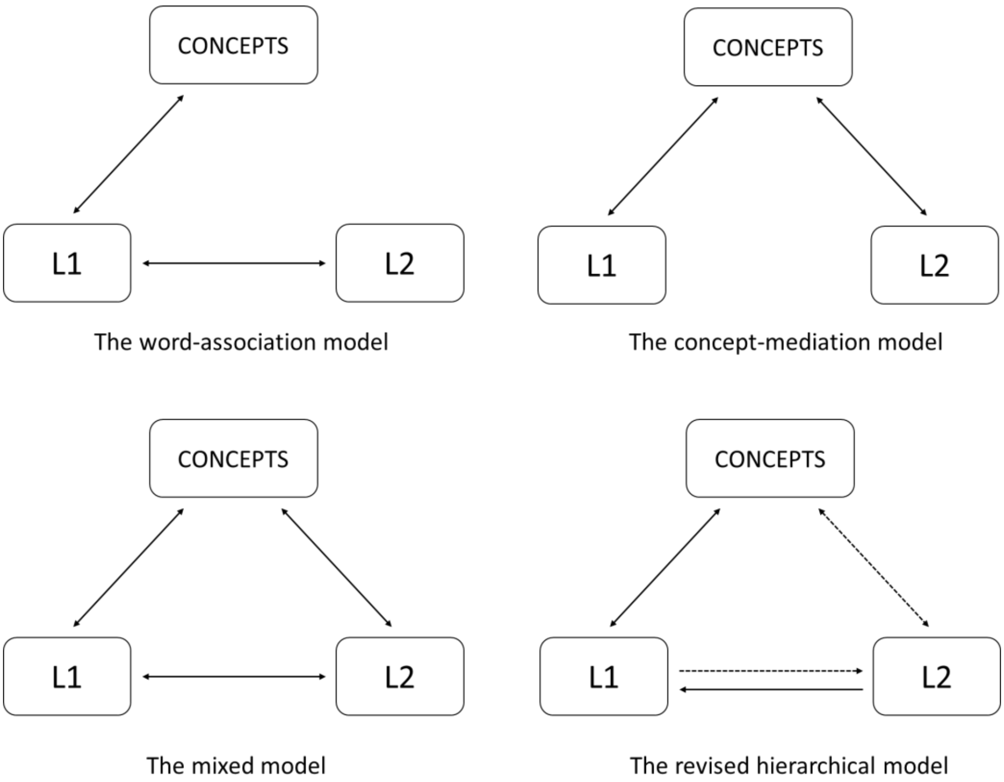


Figure 2.1. Four types of bilingual lexicon models. Adapted from French and Jacquet (2004)

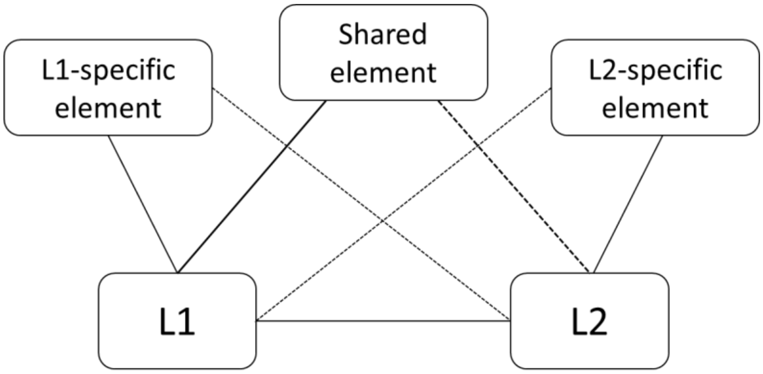


Figure 2.2. The shared asymmetrical model. Adapted from Dong et al. (2005)

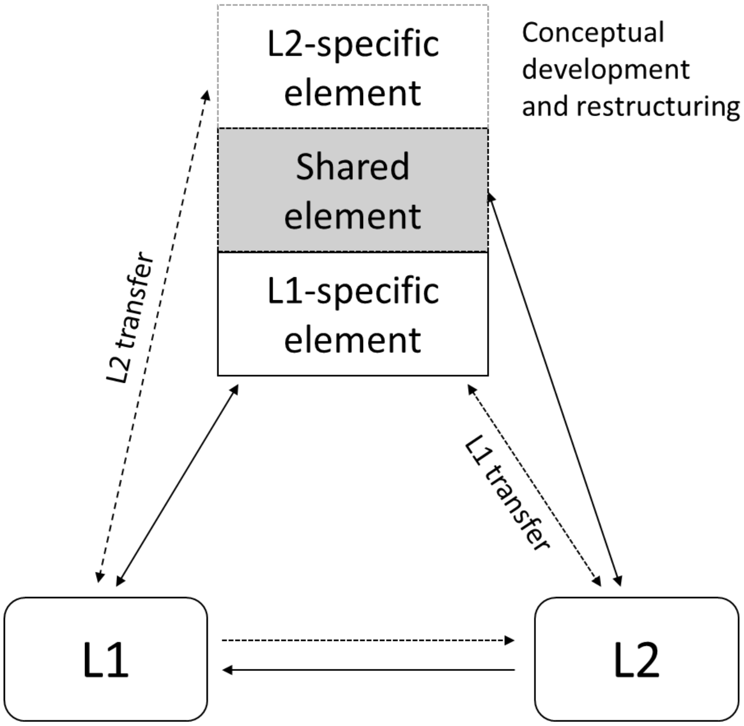


Figure 2.3. the modified hierarchical model. Adapted from Pavlenko (2009)

L2 embodiment processing can be predicted based on L2 mental lexicon models. Prediction will depend on how the lexical and conceptual links are assumed to be connected. As suggested by *the word-association model*, if L2 is linked to L1 translation equivalents ("lexical link") but not directly linked to a concept ("conceptual link"), the simplest prediction would be that L2 words may not be embodied without a direct connection to the concept. But another possibility is that they are embodied by detours through L1 and ultimately accessing the concept. However, since it is possible only through L1 translation equivalents, it can be predicted that the embodiment will be slower or impoverished. In fact, previous studies have demonstrated that embodiment processing occurs only with a brief contact of a novel word and the embodiment of L2 is weaker than that of L1. As proposed in *the concept-mediation model*, if L1 and L2 are identically connected to the concept, both L1 and L2 will be embodied and there will be no difference in the degree of the embodiment, particularly for learners with high proficiency. In fact, several previous studies found the same degree of embodiment in L1 and L2, which is mediated by proficiency. According to the assumption of the *revised hierarchical model* that the connection between L2 and the concept is relatively weak compared to that of L1, it is also expected that weakened embodiment processing will be found in L2 learners. Meanwhile, since *the shared asymmetrical model* or *the modified hierarchical model* assumes that L1 and L2, each can have its own unique connection for the same concept, the embodiment may be different depending on the concept's characteristic. For instance, abstract concepts, due to greater heterogeneity and influence of cultural differences, may have different embodiments for different languages.

2.2.3. L2 Learning environment and knowledge types

It is considered that the environment in which a language was learned has a significant impact on the formation of its embodiment. Since embodiment theory not only assumes a learning principle of language similar to that of the Emergentist approach of SLA ("learning from experience with external input") but also it emphasizes the mind-body-world interaction rather than linguistic association, consideration of the learning environment will be more crucial for the theory. This chapter will illustrate the learning environment of L2 learners, especially who learn L2 as a foreign language, with particular focus on vocabulary learning. In addition, it will attempt to predict how such a difference in environment will affect the formation of knowledge on L2 vocabulary in terms of the type of knowledge suggested in the SLA literature.

2.2.3.1. Knowledge types

The types of knowledge dealt with in the SLA literature will be described first in order to characterize the environment of L2 and what kind of knowledge that environment forms. Various attempts have been made to classify L2 knowledge. The most traditionally known is the distinction between *acquisition* and *learning* proposed by Krashen (1976). It is argued that there are two independent systems of language learning. One is *acquisition*, which is used when learners speak focusing on meaning. This is similar to how children develop the ability to use their native language. *Acquisition* is an unconscious process, in which learners are not aware that language has been acquired and they are not able to explain the rules they have acquired. They can just "feel" it and only know that they can use language to communicate. The other is *learning*, which used to confirm whether the form of that produced by the acquired system is correct. This is comparable to the way the learners develop their abilities in L2. *Learning* is the process of consciously obtaining the rules of language, so that the learners can notice what has been learned and explain its rules.

Another distinction of knowledge is *declarative* and *procedural* knowledge (Cohen et al., 1997; Ullman, 2001a; 2001b; 2014), which are acquired through declarative and procedural memory, respectively. These memory systems are considered to be the most important long-term memory systems and considered vital for language acquisition (Ullman, 2014). *Declarative* memory is an important memory system for learning information about facts or events and is responsible for storing item-based elements such as lexicons in languages. Hippocampus and medial temporal lobe are recognized to be critical for learning and storing new knowledge in this system. The regions underlying the input and output processing of new information is the triangular and orbital parts of the inferior frontal cortex (Brodmann area 45 and 47). Brodmann area 45 is part of Broca's area, which is assumed to play an important role with language production. Learning through the declarative memory system is relatively fast, and it begins to develop in childhood, stabilizes from adolescence to young adulthood, and then declines. Therefore, learners after adolescence are better at learning with this system than young children. On the other hand, *procedural* memory underlies the processing of various activities and functions, such as sequence, rule, and categorization, and is specialized in predicting probabilistic outcomes such as the next item in a sequence or the outcome of a rule. Linguistically, it is related to rule-based language processing such as processing of syntax and morphemes. Procedural memory is known to be neurologically related to basal ganglia and frontal regions. Basal ganglia are vital in acquiring new motor-cognitive skills, and frontal areas, particularly the premotor cortex (BA 6) and the opercular part of the inferior frontal cortex (BA 44), which is also part of Broca’s area, is important in the processing of automated skills. Learning through procedural memory is relatively slow and requires more practice, but the result of learning is fast and automatic processing. Compared to declarative memory, it matures in childhood, declines before adolescence, and weakens in adulthood. From the perspective of declarative/procedural memory system, it is predicted that L2 acquisition after adolescence is difficult particularly in syntactic related properties, because syntactic knowledge is closely related to procedural memory system and it declines in adulthood.

The most prevalent distinction of knowledge is *implicit* and *explicit* knowledge. Declarative knowledge and learning are regarded as a form of explicit knowledge, and procedural memory and acquisition as implicit knowledge. However, they are not completely identical concepts. The difference is that "awareness" in the learning process plays an important role in the distinction between implicit and explicit knowledge (Hulstijn, 2005). Explicit knowledge is obtained through explicit learning, which means conscious learning of what is being learned. Conversely, implicit learning is learning without awareness of what is being learned, occurring unconsciously. Declarative/procedural memory is not affected by the presence of consciousness as it is based on systemic differences in the brain. Knowledge by declarative memory system can be called declarative knowledge, and it is generally assumed to be acquired through explicit learning, but acquisition through implicit learning cannot be excluded (Ullman, 2014). Of the relevant questions regarding the types of knowledge in the field of SLA is the interface between explicit and implicit knowledge. Researchers have questioned whether explicit knowledge can be converted to implicit one, or whether explicit knowledge aids in the formation of implicit knowledge, or whether the two knowledges are completely separate systems (VanPatten & Williams, 2015a).

According to Krashen's monitor theory, because acquisition and learning are completely separate, explicit knowledge cannot become implicit knowledge (VanPatten & Williams, 2015b). On the other hand, usage-based approach assumes that the explicit knowledge can be turned into implicit knowledge or knowledge that are functionally equivalent to it, through practice, proceduralization, and automatization (Ellis & Wulff, 2015; DeKeyser, 2015).

2.2.3.2. L2 Classroom setting

L2 learners encounter L2 mainly in the classroom situation. Despite the advancement of technology and increased access to the internet and mobile devices, which have the potential to create opportunities for L2 usage in daily life (Kacetl & Klímová, 2019), many learners still face challenges in using L2 outside of the classroom. The classroom environment differs from natural settings in terms of the quantity, quality, and type of input available for language acquisition (Gass et al., 2020). First, the quantity of input available is relatively limited. There are three primary sources of input in this context: instructors, textbooks, and other students. The largest source of input is the one from the instructors, but since they typically have a large number of students, the input each student receives from the instructor is minimal, and opportunities for interaction with the instructor are even more limited. While immersion education and new mobile technologies may provide increased access to the target language outside of the classroom, it is still less than that which would be available in a natural setting, where learners are able to constantly and inevitably receive input from everyday life. The input is not only limited in terms of quantity but also in quality. Above all, interaction, a form of inputs that can play a significant role in language learning, is particularly lacking in a classroom situation compared to the natural setting. Moreover, most inputs are provided with a restricted or modified form, often called foreign-talk, analogous to baby-talk. It is known that the complexity of the syntax used by teachers becomes less complicated and the length of the sentence becomes shorter, when the students are less proficient (Gass et al., 2020). This has the advantage of providing an input suitable for students, but also reduces the chances for students to encounter different levels of language. L1 children are also exposed to modified language, but since parents or instructors are not the only source of their input, they can come into contact with more diverse forms of language. Furthermore, teachers have relatively limited knowledge of L2 because they themselves are L2 learners in most cases. The interaction with instructors, which is significantly lacking in classroom settings, is often complemented by the activities with other students. However, because other students have the same inadequate language knowledge, the input from them is of poor quality. Often, interactions between students are accompanied by errors, and they may proceed without recognizing these errors (Gass et al., 2020).

In general, the type of knowledge obtained in L2 environment is more likely to be explicit knowledge, in contrast to the natural environment, in which learning is primarily implicit (Ellis & Wulff, 2015). First of all, implicit learning is unlikely because L2 learners experience L2 relatively less in everyday life. In addition, in L2 classroom settings, explicit learning is preferred, where learners can consciously focus on the form-meaning association, because it is critical to provide a various type of L2 knowledge to the students in a timely manner (Newton, 2019). Often, the input is modified in a way that makes it more prominent to the learners. This is called input enhancement, which is intended to draw the learners’ attention to target forms, allowing them to notice the form of the input (Kormos, 2019). In much of the SLA literature, *noticing* is regarded to play a crucial role in learning L2 (Alcón, 2007). However, not all L2 vocabulary knowledge consists of explicit knowledge. Explicit knowledge can be converted into implicit knowledge through practice, and the changing L2 learning environment provides opportunities to implicitly encounter L2 in everyday life through various media, which includes implicit learning through context surrounding words in text. However, since these opportunities are still relatively scarce, it is likely that explicit knowledge will prevail in the way that knowledge of L2 is represented.

2.2.3.3. Vocabulary learning in classroom setting

It is extremely difficult to teach vocabulary particularly in a timetabled classroom setting (Newton, 2019). The main issue is the enormous number of words that must be taught in a short amount of time. It is assumed that 2000-3000 words are essential for conversation and 8000-9000 words are needed to read academic texts (Nation, 2013b). However, the number of words that can be achieved during class seems to be too small compared to that. For example, according to Webb and Chang (2012), Taiwanese L2 learners who receive instruction for 2 to 4 hours per week achieved only 18-282 words per year. Furthermore, not only size but also depth are key factors in word acquisition. Knowing vocabulary includes knowledge of collocation and word usage. Both word breadth and depth are known to be highly connected to L2 proficiency (Qian & Lin, 2019; Yanagisawa & Webb, 2019). In short, it takes much time and effort to learn vocabulary, but the problem is that it is not possible to devote so much time to vocabulary learning in a classroom situation.

Intentional learning is known to be necessary in the initial stage to learn about 2000-3000 words required at the beginning of L2 learning (Kormos, 2019). This type of learning involves a cognitive process of learning as a goal rather than resulting in an incidental outcome and requires a high level of attention (Webb, 2019a). Its success depends primarily on learners' motivation and is often carried out by explicit, language-focused instruction. The main results of such education are explicit knowledge (Webb, 2019a), but it is also presumed that they can be converted into implicit knowledge through practice. In general, it is a concept close to, but not identical to, explicit learning. Because intentionality is important in this term, even explicit learning can be regarded as incidental learning when intentionality is insufficient (Kormos, 2019).

The most common type of intentional learning is to memorize a word list. The word list is mainly decontextualized so that the meaning is provided as the L1 translation equivalent, and in other cases, the context is also given in a way the word is included in the sentence (Kormos, 2019). The decontextualized word list is known to be effective in improving the vocabulary size of L2 learners (Laufer, 2006). No need to process context is rather helpful because it makes students pay more attention on the form-meaning association (Qian, 1999). In contrast to the development of vocabulary breadth, it is questionable whether vocabulary depth is improved as well by explicit learning from such a word list. However, some studies found that the depth of word knowledge acquired through a decontextualized word list was not significantly different from that from a contextualized word list (Webb, 2007). It seems that L2 learners successfully apply the word-related knowledge of L1 to L2 vocabulary through inferring process (Webb, 2007).

On the other hand, incidental learning is learning words without intention, usually meaning learning words that are included in text but not focused (Webb, 2019b). Many studies emphasize the importance of being exposed to extensive and intensive reading and listening texts when developing vocabulary through incidental learning (Brown et al., 2008). This is based on the findings that a large amount of input is beneficial for children's vocabulary development in the study of L1 acquisition. About 3000 basic-level words can be learned mainly through incidental learning, and deeper knowledge of vocabulary is in most cases acquired through incidental learning (Kormos, 2019). However, it is unclear how effective this method of learning is for L2 learners (Vidal, 2011). Learners frequently ignore unfamiliar words and skip over them, especially if those words are not salient, or not required to understand the whole context (Vidal, 2011). Also, even if the learner attempts to guess unknown words, successful reasoning strategies cannot be used unless a sufficient level of lexical knowledge is attained (Nation, 2013a). Furthermore, for a word to be acquired by incidental learning, it must be encountered repeatedly in different settings (Crossley et al., 2013). It is generally assumed that repeated exposures of 5 to 16 times must occur, and these encounters must take place before previous ones are forgotten (Webb, 2007). Since in a typical L2 environment, repetitive encounters satisfying these conditions are difficult, it seems challenging to learn a large vocabulary in the initial stage by relying only on incidental learning.

Since the knowledge distinction in SLA does not generally focus on vocabulary, it may be difficult to classify vocabulary knowledge according to the above classification of knowledge type. Krashen (1976) divided the types of knowledge into the learning/acquisition dichotomy and posited that L2 learners primarily rely on learning instead of acquisition, but this focuses more on speech and grammatical analysis rather than lexical knowledge. Since the vocabulary learning of L2 learners is both intentional and incidental, the form of L2 vocabulary knowledge can be described either with acquisition or learning (or implicit or explicit). Given the difficulty of incidental learning in early vocabulary learning, it is predicted that L2 vocabulary is more likely to be explicitly learned, not implicitly acquired at least in the early stages. However, it seems less persuasive to argue that the "acquisition" of L2 vocabulary is impossible, since it is clearly possible, although less effective, for L2 learners to acquire vocabulary through incidental learning as shown in the previous studies (Brown et al., 2008; Crossley et al., 2013; Vidal, 2011).

As for the distinction between implicit and explicit knowledge, because an interface between the two knowledge types is possible (Ellis & Wulff, 2015), this issue can be resolved more easily. Even vocabulary knowledge obtained through explicit learning can be converted into implicit knowledge through repeated encounters and practice. In this regard, it is predicted that the vocabulary knowledge of L2 is mainly close to explicit knowledge in the initial stage but becomes implicit knowledge as they repeatedly have more input.

Meanwhile, from the perspective of the declarative/procedural memory system, lexical items are considered to be learned by the declarative memory. Since this is the same for both L1 and L2, it is predicted that there is no difference in the type of lexical knowledge prevailing in L1 and L2.

2.2.3.4. Embodiment and knowledge types

Provided that embodiment is also a type of knowledge, it is possible to classify embodiment relating it as a type of knowledge. Although embodiment theory does not discuss it in terms of types of knowledge, considering the principle of embodiment, it can be predicted that it has a closer relationship with implicit knowledge. The association between language and perception, action, and emotion forms an embodiment, and these associations are not consciously connected by the learners, but rather occur through daily experiences. However, as shown in previous studies using novel word paradigm, the conscious association of gestures and language also formed embodiment. In the process of language acquisition, it is often learned by directly referring to an object, and this connection between the referent and language is rather closer to explicit learning. For this reason, explicit knowledge and embodiment are not incompatible. However, since such conscious association is not so common except when encountering a certain concept for the first time, it is assumed that embodiment is, in general, more closely related to implicit knowledge.

In this respect, the environment of L2 vocabulary learning is less suitable to form an embodiment. This is first because explicit learning occupies a substantial amount of L2 learning. Moreover, the associations that L2 learners face in the L2 environment are to some degree different from those in the L1 environment. In L2 vocabulary learning, since explicit learning mainly consists of learning through translation equivalents, an indirect connection through language is more probable rather than a direct connection between language and referents. Implicit learning is also more likely to be achieved through an indirect way such as reading or listening rather than interaction with the actual environment surrounding the language compared to L1. Therefore, it can be presumed that the embodiment of L2 occurs through explicit and indirect implicit learning mediated through language. Exploring the embodiment phenomenon of L2 can be helpful in studying how this indirect form of learning might affect the embodiment.

The biological basis for declarative/procedural memory is primarily tied to linguistic areas like Broca's area, which is more relevant with amodal theory. This system, however, does not rule out the possibility of an embodiment. The premotor cortex, which is associated to the motor domain and in which the activity of mirror neurons is found, is one of the key neurological foundations of the procedural memory system (Ullman, 2004). Mirror neurons are known to be activated during observation and reproduction of action (Rizzolatti & Craighero, 2004). Additionally, procedural memory and embodiment share certain theoretical similarities, in that procedural memory is the memory of motor and cognitive abilities and habits, whereas embodiment is the re-enactment of sensory-motor experiences. In this respect, the embodiment appears to be related to the procedural memory system.

However, the declarative-procedural memory system is conflicting with embodiment cognition in that it separates the memory system by the properties of language. According to the declarative-procedural memory system, learning lexical objects is associated with declarative memory, while learning grammatical rules is associated with procedural memory. Embodiment cognition is incompatible with this distinction, which claims that every aspect from lexical items to grammatical rules is grounded in the sensory-motor areas. Particularly, it is contradictory that vocabulary learning occurs only through declarative memory. From the perspective of embodiment cognition, understanding vocabulary does not seem to be simply a process of input and output of individual items, but rather a total representation of the experience surrounding that item. In this regard, embodiment cognition and the declarative-procedural memory system appear to be incompatible. However, both have their own reasonable evidence, so there is no need to discard one of the theories and exclusively choose the other. Based on the evidence of each, it is possible to seek a way to reconcile two theories by making modifications to each theory and addressing questions such as whether lexical items can be processed in procedural memory systems and why linguistic-specific areas are activated during language processing.

2.3. General research questions of the dissertation

This section presents the general research questions of this dissertation based on the literature review outlined above. This dissertation conducts two experiments in order to answer the following three questions. (i) Is the embodiment of lexical processing essential for comprehension or mere by-product from amodal semantic processing? (ii) Are abstract concepts embodied as well? If so, on what are they grounded? (iii) Is L2 processing embodied as well? If so, how is it distinguished from L1 processing? Each question is described in depth below, along with the experiments that will be undertaken to address it.

The first research question of this dissertation is whether embodiment is essential for understanding language. From the perspective of embodiment cognition, a number of studies have been conducted on whether embodiment is crucial for language processing, that is, whether modality-specific areas related to each sensory-motor modality are engaged during language processing. Much of research has demonstrated activation of modality-specific areas such as the motor-cortex during language processing, which is considered evidence supporting the embodiment cognition approach. One problem with this claim, however, is that the activation of a certain brain area during a certain activity does not establish that the region is indispensable. Instead, it may be a mere by-product, or result of the activity of other amodal areas specialized for amodal semantic processing. To address this question, this dissertation conducts an EEG experiment (Experiment 2) and measures whether modality-specific neural responses occur during early, restricted lexical processing, using the masked prime repetition technique. If the areas corresponding to each sense are activated during language processing at the limited lexical processing from the early time-window, it would suggest that embodiment is not a mere by-product, but may be necessary for language processing.

The second research question is whether abstract concepts are also embodied in language processing, and if so, into what they are grounded. Since the majority of embodied cognition research focuses on action verbs and related activation in motor cortex, it is necessary to conduct the research on other modalities or features beyond action verbs. Among them, the greatest challenge is the embodiment of abstract concepts, which seem to have no tangible referent in the world to ground. Without an adequate explanation of the embodiment of abstract concepts, the embodiment cognition theory will become an incomplete theory that fails to account for greater portion of language. This dissertation seeks to explore whether the embodiment of abstract concepts can be also explained with the activation of modality-specific areas, and if so, what modality they are related with. This study paid attention, as a sensory modality to understand the embodiment of abstract concepts, to interoception, which is known to have a correlation with abstractness. In Experiment 1, it will be investigated whether participants employ interoception as a sensory modality to process words, whether it is distinguished from other senses, and whether it is necessary for language processing in terms of embodiment. Experiment 2 will examine through EEG experiments whether the sense called interoception is actually distinguished from other senses in neural response and whether there is a neuroanatomical ground for it.

The third research question is whether L2 processing is also embodied and how it differs from L1 processing. Previous research on L2 processing has indicated that it may be quantitatively or qualitatively different from L1 processing, yet there have been relatively few studies that have examined the embodiment of L1 and L2 processing. Given that the basic assumption of embodiment cognition is predicated on the interaction between language and the environment, it is possible that the environmental differences between L1 and L2 may result in significant variations in the degree of embodiment. In order to answer this question, Experiment 1 investigates whether L2 speakers are able to recreate sensory experiences through a minimal contact with words as L1 speakers, and Experiment 2 will examine if the speaker's capacity of associating words with sensory experience is actually reflected in neural responses and how this differs from that of the L1 speaker.

Chapter 3. Neurological background

This study utilizes EEG (Electroencephalography) as the main methodology to explore the embodiment processing of L1 and L2 users. This chapter describes the functional background of EEG and explains the procedure of several EEG analysis techniques including Event-Related Potentials (ERP), time-frequency analysis and source estimation analysis.

3.1. Electroencephalogram (EEG)

EEG is a neurophysiological methodology that measures and records brain waves or electrical signals generated in the brain. Compared to other methods such as Functional Magnetic Resonance Imaging (fMRI), Positron Emission Tomography (PET), and Single Photon Emission Computed Tomography (SPECT), it has advantages in non-invasiveness and high temporal resolution (Lystad & Pollard, 2009; Figure 3.1). Since EEG is measured through metal electrodes placed along the scalp, neural responses can be obtained without emitting radiation and affecting participants. In addition, it has a high temporal resolution of one thousandth of a second, it can provide information on brain activity related to human cognitive processing that occurs within a second. Due to these advantages, it used to be employed for clinical purpose of diagnosing neurological disorders such as epilepsy or distinguishing states of alertness, drowsiness, and sleepiness, and now it has been widely used for the purpose of exploring human cognitive activity such as language processing (Gasser & Molinari, 1996).

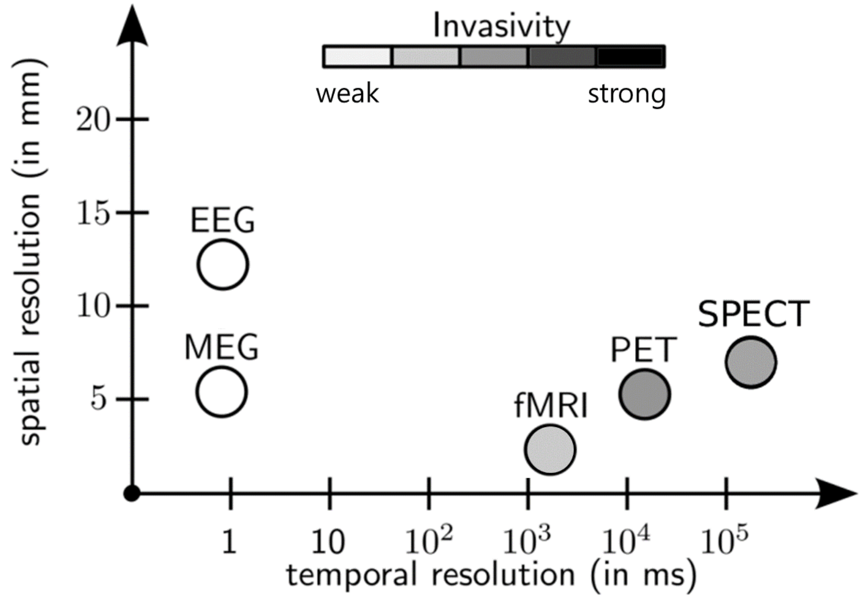


Figure 3.1. Comparison of spatial and temporal resolutions and invasiveness among neuroimaging techniques. Adapted from Lystad & Pollard (2009)

Electric signals measured through EEG reflect the activity of neurons located in or near the cerebral cortex under each electrode. During the activation and interaction of the neurons, two main types of electrical signals are generated: action potentials and postsynaptic potentials (Luck, 2014). Among these, postsynaptic potentials are, in most cases, the one obtained during EEG recording. An action potential is an electrical signal occurring within a single neuron, where it transmits information from a dendrite, or the receiving part of a neuron, to an axon, or the transmitting part of the neuron. It plays a fundamental role in communication between neurons. Its potential is relatively large, rising from its resting potential level of -70 mV to around +50 mV, but it lasts for only a few milliseconds (1 to 10 msec). So, it is easy to isolate an action potential from a single neuron but it cannot be detected by scalp electrodes due to its short duration. On the other hand, a postsynaptic potential is the one generated when neurons interact with each other at a synapse. A synapse is a gap between neurons, where a neuron transmits information to another neuron through chemical signals called neurotransmitters. When a neurotransmitter binds to a postsynaptic neuron, an electrical signal is generated, which is a post-synaptic potential. Postsynaptic potentials are relatively weak (0.1 to 10mV) but last for a longer period of time (5 to 100ms). Because of their relatively long-lasting nature, they can be summed over time and across space in contrast to action potentials. That post-synaptic potentials can be summated is significant for EEG measurement. The potential from a single neuron is so small that a scalp electrode is not able to capture it. The summed potentials from multiple neurons, however, allow the voltage to be measured at scalp electrodes. Moreover, since the neurons of the cerebral cortex are arranged perpendicular to the cortical surface, post-synaptic potentials activated in the neuron layer of the cerebral cortex do not cancel each other but overlap, transmitted perpendicularly to the scalp. In short, neural signals collected through EEG are most likely to be the summated post-synaptic potentials from the activity of neurons at the cerebral cortex.

Although the electrical signals from post-synaptic potentials can be obtained by surface electrodes, they are not strong enough to easily pass through the skull with the high electrical resistance. Therefore, they should be amplified through an amplifier during recording in order to be meaningfully interpreted. However, even with this process, signals generated in subcortical regions such as the hippocampus, basal ganglia, and limbic system almost disappear, which suggests that most of the information collected through EEG is likely to be limited to the cerebral cortex. Furthermore, the high resistance of the skull leads to spatial blurring. For this reason, the spatial resolution of the EEG is relatively low in contrast to its high temporal resolution, meaning that it is only possible to roughly estimate the source of the signals. However, in recent years, as a variety of source analysis techniques have been developed, researchers are increasingly applying source estimation through EEG (Jatoi & Kamel, 2017).

3.2. Event-Related Potentials (ERP)

ERP (Event-Related Potentials) technique is one of the methods to extract the neural responses for a specific cognitive activity. The specific responses can be obtained by segmenting the EEG responses based on the time point at which a certain event occurs and then averaging them. ERP is the term referring to neural responses extracted by this way. The term, *Event-Related Potential*, itself implies that these responses are the electrical signals segmented based on a specific event. The averaged neural response to a specific event is represented as an ERP waveform such as Figure 3.2.

Even though the details of ERP analysis may vary depending on the type of event, individual ERP waveforms are generally calculated as follows, First, researchers measure the brain waves of participants through EEG device ("*recording*"). The most commonly used method for attaching EEG electrodes is the 10-20 system (Figure 3.3). The electrodes are placed in regions of the brain divided into pre-frontal, frontal, central, parietal, temporal, and occipital. They are placed at intervals of 10 or 20 percent from the front nasion to the rear inion, centered on the calvaria area (Cz), from which the name 10-20 system is derived. While measuring participants' EEG, they were presented with specific stimuli, or events. The time point when the stimulus is presented and what kind of stimulus is presented at that time are recorded along with the EEG. This process is called *triggering*.

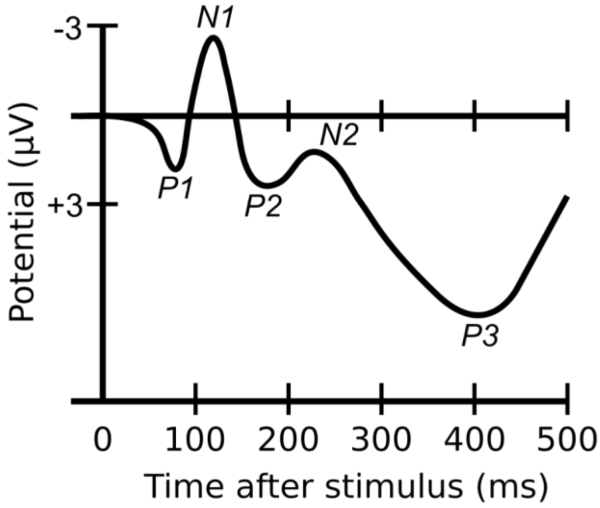


Figure 3.2. An example of ERP waveforms and ERP components. Negativity is plotted up which is the plotting convention of the early days of ERP research (Luck, 2014). However, in the remainder of the paper, waveforms were plotted with positive upward following the mathematical convention for the convenience of the general readership.

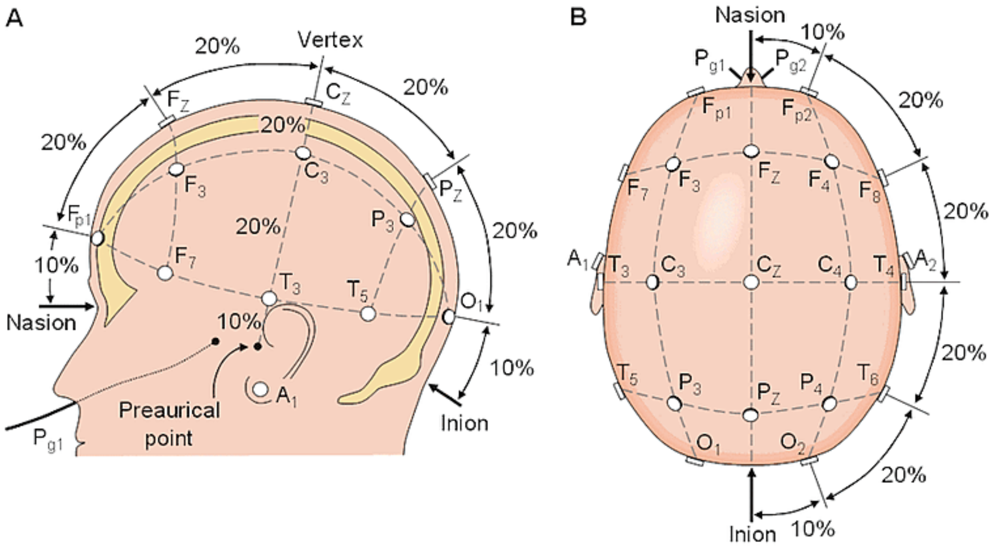


Figure 3.3. International 10-20 system for electrodes placement. Fp= frontal polar, *F=frontal, C=central, P=parietal, T=temporal, O=occipital, A=ear lobe*. Adapted from Shriram, Sundhararajan & Daimiwal (2013).

In most cases, the collected raw EEG signals should be filtered and corrected in order to eliminate noise caused by the muscle movement, eye blinking, power line, and interference with other devices. *Filtering* is to attenuate a specific frequency band to reduce noise. For example, since the electromyogram (EMG) is mainly composed of frequencies above 100 Hz, signals caused by muscle activity can be reduced by suppressing them (Luck, 2014). *Artifact rejection* is to exclude part or all of epochs that clearly contain noise from analysis. For instance, an epoch that includes eye-blink forms a large deflection as shown in Figure 3.4. Since epochs containing this type of noises have a substantial impact on the overall data, noise need to be reduced by removing these epochs. However, removing an entire epoch to reduce noise may rather lower the signal-to-noise ratio by decreasing the total number of epochs included in the analysis. Thus, if possible, it is better to subtract away the noises from the data by applying *artifact correction*. Numerous techniques for artifact correction have been developed, which may be broadly classified into three categories: regression-based approaches, dipole localization approaches and statistical component isolation approaches (Luck, 2014). Among them, Independent Component Analysis (ICA) is one of the most widely used statistical approaches, which will be employed as artifact correction method in this dissertation as well. ICA refers to an analysis method that statistically decomposes mixed signals. ICA-based artifact correction is to apply this statistical method to the EEG to subtract the noise away by unmixing the raw data. Figure 3.4 depicts waveforms in which blinks have been eliminated after applying ICA-based artifact correction. After filtering and artifact rejection/correction, the EEG responses are segmented based on specific time points recorded during triggering ("segmentation"). Then, baseline correction is performed, where segmented waveforms are moved upward or downward in order to align different amplitudes at the onset of each segment into zero. By averaging these segments, an individual ERP waveform for a specific stimulus is calculated. If necessary, the individual ERP responses are averaged to obtain a grand-averaged ERP waveform for a specific stimulus. This process of converting raw data into a form that can be analyzed is called pre-processing (Figure 3.5).

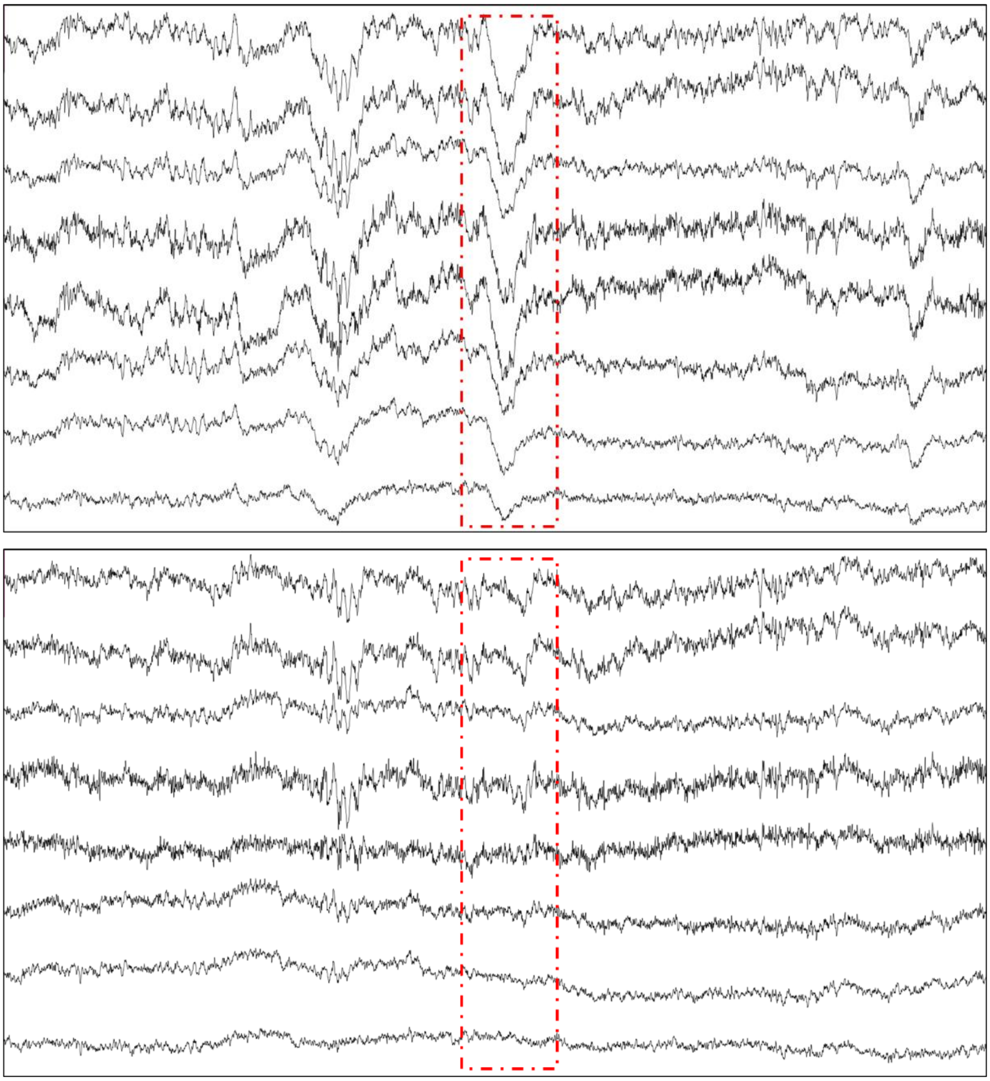


Figure 3.4. Ocular correction. [Top] An example of eye blinks. Large deflections are noticeable in the red square. Epochs with eye blinks can be either rejected or corrected. [Bottom] An example of the corrected epoch with eye blinks. After correcting the noise, the large deflection by eye blinks is minimized.

The ERP waveform shows the average change in EEG after an event occurs. As seen in Figure 3.2, there are several peaks on the waveform. Based on these peaks, the ERP waveform can be classified into several ERP components. Each component is generally named according to the characteristics of those peaks such as the polarity and peak time. For example, N1 represents the negative ERP component peaking at about 100 msec after the onset (or the first negative peak), while P2 or P200 represents the positive component around 200 msec after the onset (or the second positive peak). It is suggested that each ERP component reflects a different cognitive processing. For example, N400 and P600, the two most widely known components in language processing, are considered as indexing semantic violations and syntactic anomalies, respectively (Kaan et al., 2000; Kutas & Federmeier, 2011). A detailed description of several ERP components related to the experiment in this dissertation is provided below.

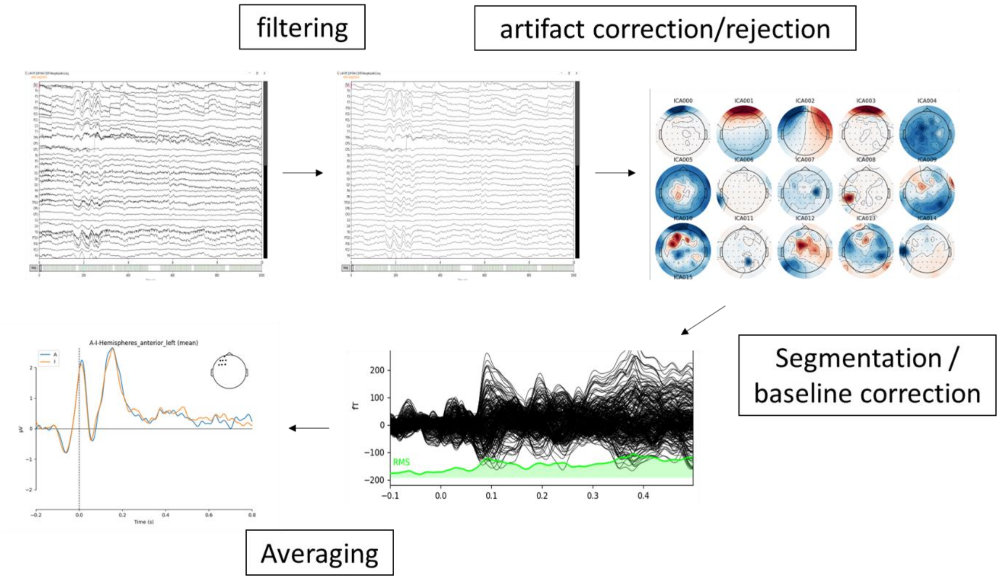


Figure 3.5. An example of pre-processing for ERP analysis (filtering, artifact correction/rejection, segmentation/baseline correction, and averaging)

3.2.1. ERP components

P2 (N1-P2 Complex)

P2 (or P200) is a positive going component that peaks between 150 and 250 msec after the stimulus onset (Crowley & Colrain, 2004). Because it usually follows the first negative polarity, N1, the two components are often collectively referred to as the N1-P2 complex. P2 is one of the components evoked in early perceptual processing as part of a normal response to visual (or auditory) stimuli (Hillyard et al., 1998). Although its amplitude and latency are most likely to be affected by physical properties related to sensory perception, such as size, shape or color of stimuli (Luck & Hillyard, 1994), they are also modulated by various cognitive factors such as attention or working memory (Dunn et al., 1998). For example, repeated visual stimuli evoked a larger P2 amplitude when attended to (Hillyard et al., 1998; Misra & Holcomb, 2003). Therefore, P2 may not only reflect early perceptual processing but also other cognitive activities of that time period.

In language processing, P2 is mainly influenced by perceptual features of written words, such as orthography or length (Hsu et al., 2009). From a traditional point of view on word processing, it is considered that the 150–200 msec after a word presentation is primarily devoted to the extraction of perceptual features of words, while semantic features are not processed (Posner & Abdullaev, 1999). The processing related to semantics was regarded to be reflected in the neural responses after 400 msec from the onset, represented by N400 (Kissler et al., 2006). However, numerous investigations have also demonstrated that the amplitude and latency of P2 are altered by factors other than the physical properties of words, including statistical characteristics such as frequency and cloze probability (Lee et al., 2012; Penolazzi et al., 2007), linguistic properties such as word category (noun or verb)

(Preissl et al., 1995) and emotionality (Kanske & Kotz, 2007). For example, words with high emotional valence, whether negative or positive, induced a larger P2 amplitude than neutral words (Kanske & Kotz, 2007).

P3 (P300)

P3 or P300 is a positive wave that peaks around at 250 to 500 msec (but typically at 300 msec as implied by the name) maximally recorded from the midline, and often sustains until later time-windows even after the peak so that it used to be referred to as late positive components (Luck, 2014; Picton, 1992). A typical example of P3 is the neural responses induced by the oddball paradigm where low-probability items (“odd”) are presented alongside high-probability items (“standard”). Low-probability targets typically elicit a more positive voltage compared to the high-probability non-targets (Picton, 1992). Theoretically, it is explained as an individual’s response to stimuli and the evaluation or categorization of them. When a sensory stimulus is received and attended to, the brain compares the input to the representation of the previous event. If the input attribute does not change, it retains the context of the previous event; when a change is detected, the context is updated. In the process of updating, an attentional resource is allocated, which elicits a more positive neural response (Polich, 2007). Because this attentional allocation is mediated by the overall arousal level of processing system, the amplitude and peak latency of P3 may vary depending on the type of cognitive processing required by the task. For example, if the task is demanding (a high-arousal state), the task itself will require a great deal of attentional resources, paying less attention to context updating, which will result in smaller amplitudes and longer peak latency. Otherwise, it will allocate more resources to context updating, which will increase amplitude and decrease latency (Polich, 2007). Numerous studies on P3 have demonstrated that the amplitude and latency of P3 can vary depending on a variety of factors affecting the arousal level, which range from biological attributes such as age and gender (Anderer et al., 1996; Melynyte et al., 2018), to cognitive states such as motivation and emotional state (Carrillo-de-la-Peña & Cadaveira, 2000; Röschke & Wagner, 2003) and some of language-related characteristics such as word frequency and age of acquisition (Polich & Donchin, 1988; Tainturier et al., 2005).

N400

N400 is a negative-going component appearing between 200 and 600 msec after the stimulus onset and generally peaking around at 400 msec. Its topography varies according to stimuli and tasks, but in written words processing it is usually formed in centro-parietal sites, slightly larger over the right hemisphere (Kutas & Federmeier, 2011). Since N400 was first discovered by Kutas and Hillyard (1980), it has been commonly known as a component representing semantic violation of language in context, but it is now considered, as a result of extensive research conducted since then, to be related to the processing of a vast array of semantic information, which can be elicited without context (Blomberg et al., 2020) and by non-verbal stimuli such as scenes, gestures, actions, and mathematics (Kutas & Federmeier, 2011). In this respect, Kutas and Federmeier (2000) consider N400 not just a linguistic component but a comprehensive component associated with the processing of semantic memory, or long-term memory. According to Kutas and Federmeier (2000), semantic memory does not originate merely from a particular region of brain such as hippocampus, but is organized in a distributed manner over various, higher-order perceptual and sensory-motor processing areas. In fact, N400 is known to be associated with various brain regions, including language-specific areas, sensory motor areas, and several areas relevant with memory (Kutas & Federmeier, 2000; 2011).

LPC (Late Positive Component)

A late positive component or late positive complex (LPC) is a positive-going ERP component that starts around at 400 to 500 msec and continues through a later time window. It frequently appears alongside N400 and is often referred as to P3, P600, or sustained positivity/negativity (Finnigan et al., 2002), due to its similarity with them in terms of timing and polarity. LPC is associated with explicit memory recognition, which is assumed to reflect the representation of recalled information or attentional shifts toward it (Friedman & Johnson Jr, 2000; Rugg & Curran, 2007). LPC is also affected by various cognitive factors, including concreteness and emotionality. As for concreteness its amplitude is generally larger in abstract words than concrete words (Kanske & Kotz, 2007; West & Holcomb, 2000), but the opposite results are also observed depending on the task (Welcome, et al., 2011). As for emotionality, it also showed mixed results (Citron, 2012). Some studies demonstrated that emotional words evoked larger positivity (Kanske & Kotz, 2007; Schacht & Sommer, 2009), while others showed the opposite trend (Hinojosa et al., 2009) or found that the effect varies depending on whether the words are positive or negative (Kissler et al., 2009).

3.3. Time-frequency analysis

Time-frequency analysis is another way to exploit that utilizes neural signals collected through EEG. It enables the examination of EEG characteristics that are not easy to capture using ERP. This section explains the background of time-frequency analysis and the characteristics of several frequency bands of interest.

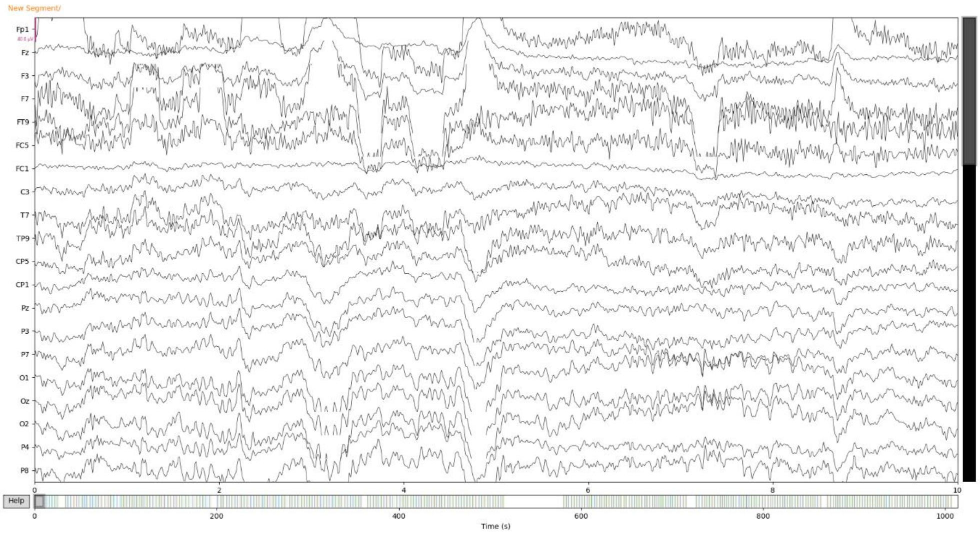


Figure 3.6. An example of EEG raw data from the participant #1. As shown in the figure, EEG signals are composed of several oscillations.

EEG signals consist of rhythmic fluctuations, as can be easily seen in raw data (Figure 3.6). These rhythmic fluctuations, or the neuronal oscillations, are known to reflect the excitability of populations of neurons (Cohen, 2014). During cognitive process, some of neurons are excited, while others inhibited by the excitation of neighboring neurons (Cohen, 2014). This synchronous firing pattern induces fluctuations in the population of neurons, facilitating information transmission (David, Kilner & Friston, 2006; Singer & Gray, 1995). These neuronal oscillations vary: some oscillations are fast, some are slow, some are long-lasting, and some are transitory, each of which indicates different patterns of neuronal synchronization. The idea is that different synchronous patterns may reflect different human cognitive activities (Başar et al., 2001; Buzsaki & Draguhn, 2004). By decomposing oscillations into several "frequency bands," certain neuronal patterns within those bands are possibly obtained, which allows one to observe specific human cognitive activities mirrored in them. These characteristics of EEG signals can be exploited to investigate ongoing human cognitive process.

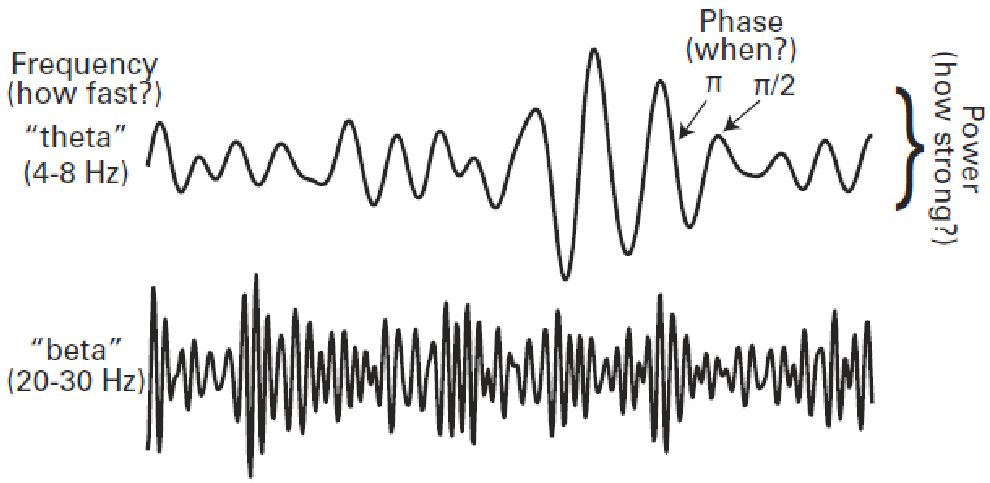


Figure 3.7. The three types of information in neural oscillations: frequency, power, and phase. Adapted from Cohen (2014).

Each oscillatory activity of EEG signals comprises three different dimensions of information: frequency, power, and phase (Cohen, 2014; Figure 3.7). Frequency describes how fast the oscillatory activity of neurons occurs, which is indicated by hertz (Hz), a unit that refers to how many cycles per second are repeated. For example, 2 Hz means two cycles per second. Power is an amount of energy, or how strong the neuronal activity is in a certain cycle, generally expressed as the square of the amplitude. Phase describes where the activity of a neuron in a specific frequency is located with respect to a sine wave at a given time point, measured in radians or degree. It is an index of whether or not neurons in a neuronal activity of a specific frequency concurrently fire. It is difficult to capture these types of information by merely analyzing EEG changes over time, such as ERP, so it is required to employ the methods such as time-frequency analysis which explores different oscillatory changes along the time axis.

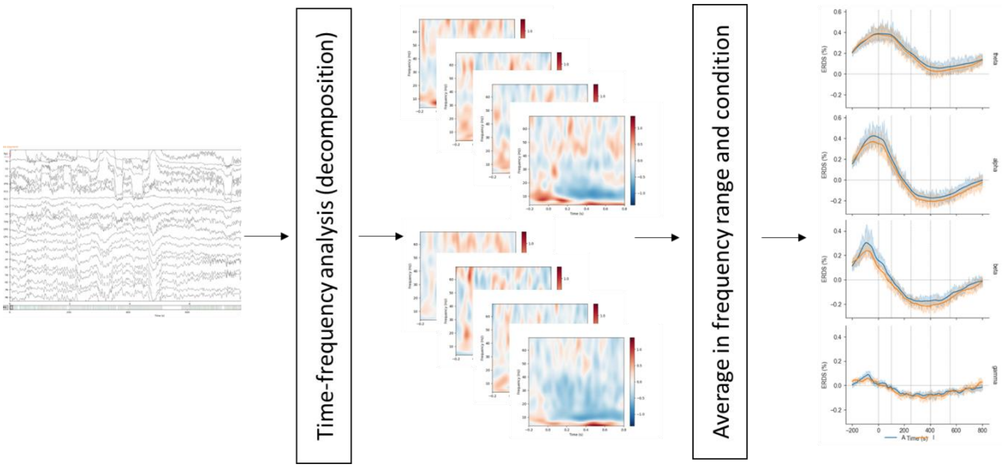


Figure 3.8. An example of time-frequency analysis process computing the total-powers (from left, EEG signals, spectrograms, and power within each frequency bands)

Although the detailed analysis method may vary, in general, time-frequency analysis is performed as follows (Figure 3.8). First, the researcher collects EEG signals from the participants. During the collection process, a particular stimulus is presented to the participants in order to examine a change in oscillation by stimuli, similar to ERP analysis. Or it is also possible to measure continuous oscillations in background activities such as sleep without presenting a stimulus. Since this dissertation focuses on neural activity induced by a specific stimulus, the process followed the method of presenting stimuli. The recording approach for EEG for time-frequency analysis is identical to that described for ERP. The obtained EEGs are pre-processed with filtering and artifact rejection and are segmented based on a designated time also similarly with ERP analysis. The difference is that a longer segment is required compared to ERP analysis since it is difficult to analyze low frequency oscillation with a short time period. The oscillation of 1 Hz frequency, for example, cannot be captured with a segment of less than 1 second since it repeats once per second. The segmented EEG signal for each stimulus is a mixture of multiple oscillations. They can be divided into each frequency by decomposing methods such as Short-time Fourier Transform (STFT), Wavelet Transform, and multitapers (Details below in the subsection, *3.3.1 Frequency domain decomposition methods*). The output of the decomposition is typically expressed as a spectrogram depicting the power change within each frequency over time (Figure 3.8. middle). The researchers calculate the average of the spectrogram of the individual stimulus based on the frequency bands of interest (e.g. 4 - 8Hz for theta band), the condition of stimulus and the relevant time window, and employ it for the time-frequency analysis (Figure 3.8. right). While averaging, to compute the time-frequency power of each trial and average all the powers produces *the total power* (decomposition → averaging). It is the most commonly used approach in the time-frequency analysis and includes both the phase-locked power and non-phase-locked power (Cohen, 2014). On the other hand, computing ERP (or averaging all trials) and decomposing it to obtain the time-frequency power generates *the phase-locked power*, or “evoked” power (averaging → decomposition). There is no clear neurophysiological basis for the distinction between phase-locked and non-phase-locked activity, but it is generally assumed that non-phase-locked activity better reflects the presence of oscillations.

Since this dissertation employs time-frequency analysis to explore neuronal oscillations that cannot be captured by ERP, total power will serve as a dependent variable.

3.3.1. Frequency domain decomposition methods

Brain wave signals recorded from EEG devices are a mixture of signals from several frequency bands. For the time-frequency analysis it is necessary to decompose this mixture. This subsection will describe several methods for decomposition, including the ones employed in this dissertation.

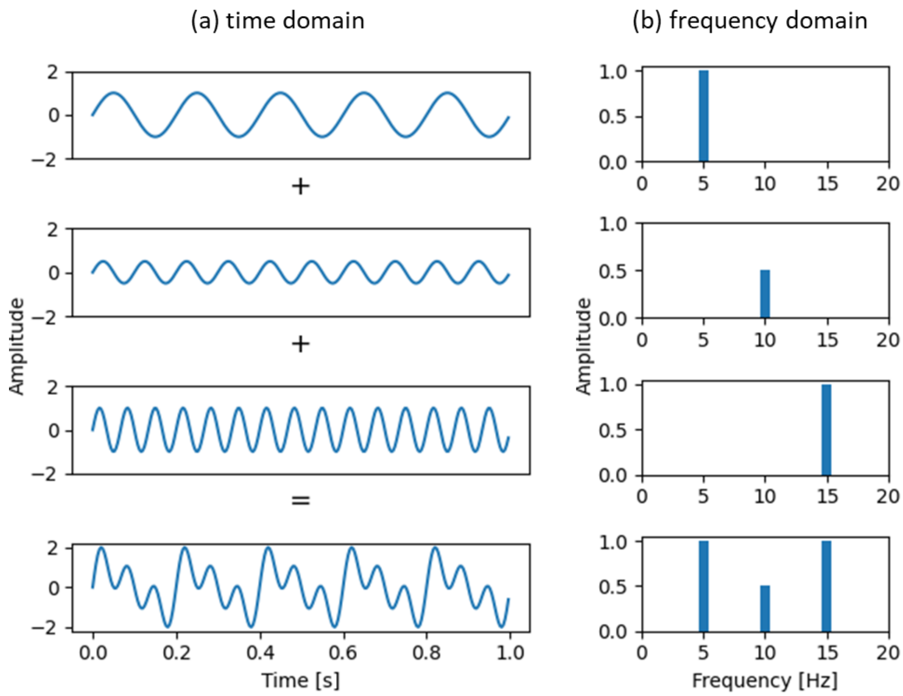


Figure 3.9. Fourier transform. [Left] Fourier transform principle: Any complex waveform can be represented as a sum of simple sine waves. Here the bottom waveform is composed of the addition of three sine waveforms with different frequencies (5 Hz, 10 Hz, 15 Hz). [Right] The result of transformation to the frequency domain. The bottom waveform was decomposed into three frequency bands.

Fourier transform is a fundamental mathematical operation of transforming time-domain data into frequency domain (Luck, 2014). Numerous different approaches for the time-frequency decomposition are in fact all based on a variant of Fourier transform (Cohen, 2014). As stated before, the waveforms of raw data (Figure 3.6) can be decomposed into several oscillations. The simplest way is to suppose that the mixture is the sum of sine waves (Figure 3.9). The French mathematician Joseph Fourier demonstrated that any complex waveform can be represented as a sum of a set of sine waves with various frequencies, amplitudes, and phases (Sueur, 2018). Fourier transform is a method of decomposing complex waveforms into sine waves.

The Fourier transform is a simple decomposition technique, but it is not applicable in all situations. In particular, it is adopted for the decomposition of continuous signals (Sueur, 2018) and all time-domain information is reduced to a single value in the transformation (Luck, 2014). Thus, alternative methods are required to analyze discrete signals while preserving time-domain information. A basic solution for this is to use a moving window Fourier analysis (Luck, 2014). In this method, Fourier transform is performed within a discrete segment rather than the entire continuous wave, and its result is computed as the power of that time window. The same procedure is carried out on successive time window, and the power at each time point is determined. However, simply performing Fourier transform within a segmented window may result in data distortion, as a sine wave is continuous, not discrete. For example, edge artifacts may occur due to sharp edges at both ends of a continuous sine wave cropped arbitrarily. Thus, an alternative window function to a sine wave is necessary, and the most widely used among them is *Morlet wavelets*. It is created by multiplying a sine wave of a certain frequency by a Gaussian function or normal distribution (Figure 3.10). Because the Gaussian function is a bell-shaped curve, the ends of the window are tapered out to zero, which reduces edge artifacts.

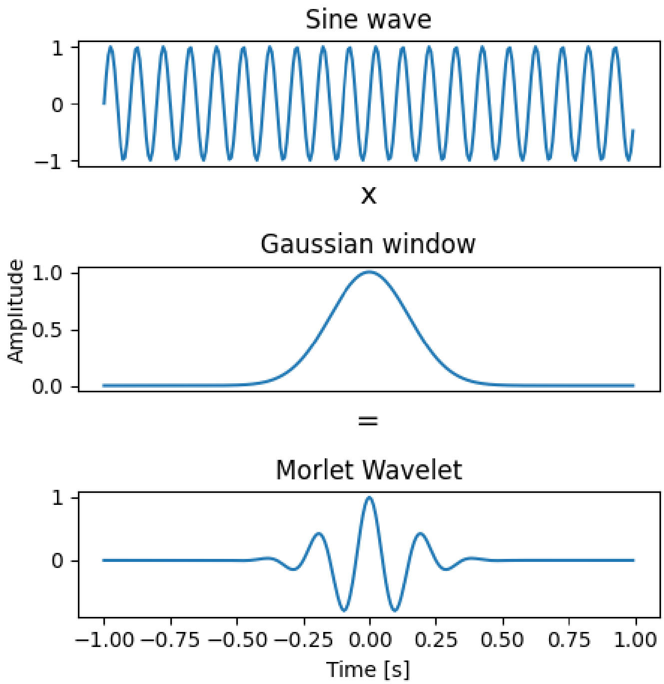


Figure 3.10. An example of Morlet wavelet (bottom), which is a sine wave (top) multiplied by a Gaussian window (middle). Adapted from Cohen (2014).

The multitapers approach is a transform method that is advantageous when the data is noisy, the number of trials is small, or the focus of the study is on high-frequency power (Cohen, 2014). It uses tapers called discrete prolate spheroidal sequences (DPSS) as a window function. In general, a single window is used to transform time-domain data into the frequency domain, but in the multitapers technique, as its name suggests, several multitapers are applied. Each power spectrum is produced by those slightly different tapers, and the average of the spectra is used as the final time-frequency power (Figure 3.11). This process of averaging spectra through multiple tapers reduces noise. However, applying this technique to lower frequency below 30 Hz is not preferred because noise smoothing through averaging interferes with frequency isolation. In the lower frequency band, isolating frequencies is relatively more important since several different brain waves from delta to beta can be identified within this range.

This dissertation will apply both Morlet wavelet and multitapers for time-frequency analysis. In high-frequency analysis, multitapers approach is used to increase the signal-to-noise ratio, whereas in lower-frequency analysis, Morlet wavelet more suitable for that range is utilized.

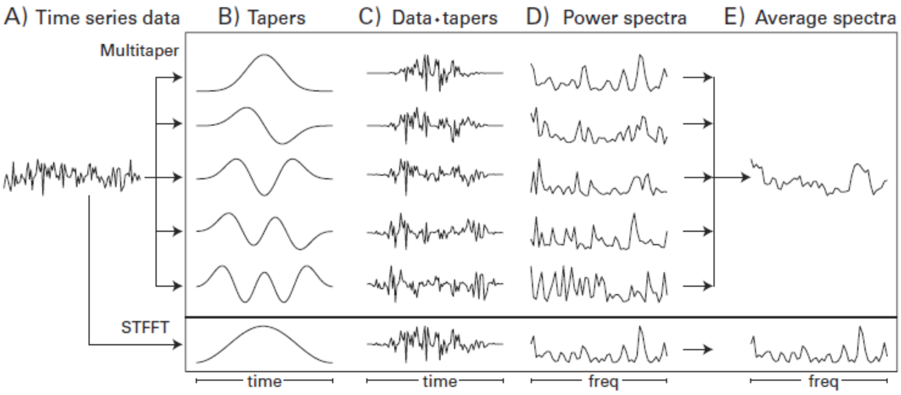


Figure 3.11. The procedure of computing power-spectra by multitapers. The short-time FFT uses a single taper, while multitpapers multiple tapers. Adapted from Cohen (2014).

3.3.2. Frequency bands and cognitive functions

As mentioned above, neuronal oscillations can be decomposed into several frequency bands, each of which is associated with a specific human cognitive activity. Although the specific range varies, EEG signals are in general divided into delta (0.5-4 Hz), theta (4-8 Hz), alpha (8-13 Hz), beta (13-30 Hz), and gamma (30-70 Hz) waves (Başar et al., 2001; Buzsaki & Draguhn, 2004; Canolty & Knight, 2010; Friederici, 2017). Following is a brief description of each frequency band in relation to human cognitive process. Delta waves are mostly found during sleep or a state of non-arousal (Hall, 2021). Theta waves are known to be related with memory formation/retrieval (Addante et al., 2011; Lega, Jacobs & Kahana, 2012; Tesche, & Karhu, 2000), while alpha rythms are associated with attention and working memory (Foxe & Snyder, 2011; Klimesch, 2012; Pfurtscheller & Da Silva, 1999; Wianda & Ross, 2019). Beta rhythms are relevant with alertness and active concentration (Gola et al., 2012; Kamiński et al., 2012) and movement and its observation (Khanna & Carmena, 2015; Pogosyan et al., 2009), whereas gamma waves are associated with extensive cognitive activities such as working memory, attention, and conscious perception (Kaiser & Lutzenberger, 2003; Miltner et al., 1999). It is also known that the neural oscillation changes of each band are related to language processing (Bastiaansen & Hagoort, 2006; Bastiaansen & Hagoort, 2015; Bastiaansen et al., 2005; Davidson & Indefrey, 2007; Hald, Bastiaansen & Hagoort, 2006; Lewis et al., 2017; Mellem, Friedman & Medvedev, 2013; Weiss & Mueller, 2012). In particular, beta and gamma bands are closely related with language processing; for example, Bastiaansen & Hagoort (2015) claimed that an increase of power within the beta band reflects syntactic unification whereas increase in the gamma band indicates semantic unification.

The frequency band most relevant to the activation of sensory-motor areas explored in this paper is gamma bands. In the last few decades, the implications of gamma bands oscillations in sensory and cognitive activities have received considerable attention in the field of neuroscience (Herrmann, Munk & Engel, 2004; Ribary, 2005). Many studies have demonstrated that gamma band is related to sensory perceptual processing including visual and auditory perception (Başar-

Eroglu et al., 1996; Pantev et al., 1993; Rodriguez et al., 1999; Steinmann et al., 2014; Strüber et al., 2000), memory (Herrmann et al.,2004; Kaiser et al., 2003; Tallon-Baudry et al., 1998), attention (Gruber et al., 1999; Tiitinen et al., 1993), and language processing (Eulitz et al., 1996; Ihara et al., 2003; Peña & Melloni, 2012; Pulvermüller et al., 1996; Pulvermüller et al., 1999). This study focuses on the gamma band oscillation because it contributes to the transmission and integration of the distributed representation of sensory information (Light et al., 2006; Tanji et al., 2005). If each different sensory information is processed in different areas of brain, one might ask how they can be integrated simultaneously to form a single piece of information. For example, in order to recognize any object, it is necessary to perceive a variety of sensory information such as its color, size, and texture and to integrate and recognize them as a single item. Neural synchronization is one of possible explanations for this, where distributed neurons interact with each other by simultaneously firing mostly in the range of gamma bands (Tallon-Baudry et al., 1997).

According to embodied cognition theory, the processing of words activates the sensory areas of the brain related to them. In other words, when a word is retrieved from memory, the regions associated with the perception of the referent of the word are activated along with the medial temporal lobe responsible for memory retrieval (Meyer et al., 2005). This co-activation produces neuron ensembles with strong connections between those areas, which will be reflected in gamma band oscillations (Pulvermüller, 1996). In fact, the previous studies showed that real word induced increased gamma band power compared to non-word (Eulitz et al., 1996; Pulvermüller et al., 1996). Moreover, such gamma oscillations may vary by the word type, depending on the connections each word type formed. For instance, nouns with connection to the visual modality and verbs with motor modality exhibited distinct powers and topography of gamma oscillations (Pulvermüller et al., 1999).

3.4. Cluster-based non-parametric permutation test

Using ERP components is not always possible in ERP analysis. Since it is essential to have a priori knowledge on EEG responses to specific events to employ a certain ERP component, without proper knowledge, analysis based on several ERP components is not appropriate. This study explores embodiment during L1 and L2 language processing, which assumes the involvement of the sensory-motor domain. So, it is first necessary to investigate ERP components related to the activation of each sensory-motor region. However, this is based on the hypothesis that language processing activates the sensory-motor area of the brain, which has often been considered controversial. Actual experimental results may contradict this assumption, and in that case, it may not be possible to fully understand the findings with only a few pre-selected ERP components or the ones previously known.

When comparing two ERP waveforms without prior knowledge, one will encounter a multiple comparisons problem, which indicates that when multiple statistical tests are run at the same time in a statistical analysis, it is more likely to produce the erroneous inference (Benjamini, 2010). Since EEG datasets consist of numerous datapoints measured at thousands of time points from tens of electrodes, numerous inferential statistical tests should be simultaneously performed to compare these numerous datapoints between two conditions. During this comparison process, significantly different datasets are bound to be discovered at certain data points. Nevertheless, it does not necessarily mean that this difference in fact reflects underlying differences of human brain activity, because erroneous inferences cannot be rejected due to a multiple comparisons problem.

The most typical and simple way to solve this problem is to strictly adjust the significance level to the number of inferential statistics being performed. For example, in the Bonferroni correction, one of the most widely known adjustment methods, the significance level is obtained by dividing a critical alpha-level of 0.05 by the number of samples. This is a simple solution to reduce the false alarm rate, but the actual effect can be easily missed because it estimates the significance level too conservatively. For example, the significance level will be 0.00005, if the number of samples is 1000, which is likely to be exceeded given the number of time points and electrodes in any ERP analysis with no existing knowledge.

Another alternative is a *cluster-based non-parametric permutation test*, which will be conducted in this study along with a traditional statistical approach or ANOVA. It is one of the specialized statistical procedures that can be employed in order to solve a multiple comparisons problem when prior ignorance on existing ERP components is taken into account (Maris & Oostenveld, 2007). This method estimates the parametric distribution based on the collected data and tests the significance of the data according to that distribution. In this process, data points are grouped into clusters and datasets are permuted by simulation so that it is referred to as a cluster-based non-parametric permutation test.

This analysis was performed as follows. (1) First, a paired-sample t-test is performed on the ERP data of two conditions at all electrodes and all time points. From the results of these t-tests, all samples larger than the t-value of the significance level are selected (significance level of 0.05). The selected samples are clustered based on temporal and spatial adjacency. Then, cluster-level t-value is calculated by summing all the t-values within each cluster. (2) Next, the cluster-level permutation distribution is calculated to confirm whether the above test statistics is a significant cluster. If the test statistics falls within 2.5% (significance level of 0.05) of both tails of the simulated permutation distribution, it is considered a significant cluster. The permutation distribution is calculated as follows. First, all trials under both conditions are collected into a single set and then randomly divided into two groups regardless of the condition. For this random partition, the cluster-level t-value is obtained as in (1). The largest of the cluster-level t-values is used as the cluster-level test statistic. This process of calculating the cluster-level statistics from random partitions is repeated a large number of times (usually 1000) and the permutation distribution is constructed through these simulations.

Non-parametric permutation test is a statistical technique used in time-frequency analysis as well as in ERP analysis. Since the time-frequency analysis additionally includes frequency as factors of statistics, the dimension of axis for the analysis is generally increased compared to ERP (ERP: time x amplitude; time-frequency: time x frequency x power). Therefore, multiple comparison problems arise more frequently with traditional statistical methods, and this is why the majority of time-frequency analysis employs a non-parametric cluster-based permutation test as a statistical tool. In this study, statistical differences between the conditions in the dimension of the time-frequency were confirmed through permutation test instead of statistical analysis based on traditional ANOVAs.

3.5. Source estimation analysis

Although EEG has been a widely used method for measuring neural activity, it has not been considered the primary technique to localize brain sources, in comparison to other neuroimaging methods such as fMRI and PET, due to its lower spatial resolution. However, the EEG localization methodology has undergone significant evolution over the past several decades, improving its spatial resolution. Empirical studies across a wide range of fields have demonstrated the reliability of these methods in localizing the cortical sources that generate the scalp recordings, often with an accuracy of approximately 15mm (Birot et al., 2014; Klamer et al., 2015; Mégevand et al., 2014; Michel, 2019). If its advantages of non-invasiveness and high temporal resolution are combined with the techniques to compensate for its disadvantage of low spatial resolution, EEG can be useful to explore the neural generators of cognitive activities processed within a few seconds.

The developments in the field of EEG source localization have been greatly influenced by the advent of distributed source imaging techniques and high-resolution EEG. Distributed source imaging techniques, which have been in use for approximately 30-40 years, are currently the most widely employed methodology for solving the inverse problem in EEG analysis. These techniques continue to evolve and improve upon the limitations of previous algorithms. This study will utilize several major algorithms based on distributed source imaging techniques to perform source estimation analysis. High-resolution EEG, also known as high-density EEG, refers to the recording of EEG using a high number of electrodes, typically more than 100. There is ongoing debate regarding the optimal number of electrodes for source analysis, but it is generally agreed that the use of more than

64 channels can help to mitigate mislocalization of affected regions (Luu et al., 2001). However, this assumption is based on the premise of minimal noise during EEG recording, which is not always achievable in practice. Therefore, the use of a high number of electrodes does not necessarily guarantee increased accuracy (Luck, 2014). In this study, EEG will be recorded using 64 channels while ensuring minimal accuracy and taking practical considerations into account.

3.5.1. Forward and inverse solutions

The process of EEG source localization involves the inverse estimation of the generators of the voltage distribution observed on the scalp. To accomplish this, it is necessary to address both the forward and inverse problem. This section will provide a definition of these problems, as well as a description of the solution.

The forward problem refers to the task of simulating the potentials measured at the scalp surface, given a known current distribution within the brain. This problem is typically formulated as a boundary value problem, in which the solution is dependent on the conductivity properties of the head and the current distribution within the brain. The forward problem is an important step in the process of EEG source localization as it provides a means of relating the neural activity within the brain to the potentials measured at the scalp. The forward problem can be solved using a variety of mathematical models, such as the boundary element method, finite element method, and finite difference method. These models rely on the use of head models which can be based on individual anatomy, population-based models, or atlases. The utilization of a head model, as opposed to the simplistic infinite homogeneous model, which assumes a uniform conductive medium throughout all space, allows for the incorporation of anatomic information and specific conductivity characteristics, such as those of the brain, skull, and scalp, into the analysis. This study employs the boundary element method (BEM), a commonly utilized technique in forward solutions, as the method of choice, which is based on the concept of solving the forward problem on the boundary of the head (Michel & He, 2019).

More complex is the inverse problem, which is the process of estimating the underlying neural activity responsible for the measured electrical potentials on the scalp. The EEG inverse problem is considered an ill-posed problem or a non-uniqueness problem, which means that there may be multiple solutions or no unique solution to the problem (Jatoi & Kamel, 2017). In the case where it is assumed that there is only one source for a certain neural activity and that source forms a single dipole that transmits electrical signals in a specific direction, solving the inverse problem is relatively straightforward. By utilizing the solution to the forward problem, it is possible to predict the voltage distribution of a model dipole and compare it to the observed surface signals. However, in reality, the data often do not have a single source, and even if they do, it is not known how many dipoles a source may have. As a result, there may be no unique solution to the inverse problem for a particular scalp distribution that is observed.

Various methods and techniques have been developed to solve the EEG inverse problem. These solutions can broadly be classified into two categories: discrete source analysis (or the equivalent current dipole method) and distributed source analysis. The former approach assumes that the spatiotemporal distribution of neural activity can be modeled with a limited number of equivalent current dipoles, which reflect the summed activity over a small cortical region (Jatoi &

Kamel, 2017). In this method, researchers place several dipoles at specific locations and orientations, and compute the distribution of voltage of these dipoles using the forward model. The predicted distribution is then compared to the actual data, and the residual variance, or error, is obtained. The position and orientation of the dipoles are then modified in order to minimize this residual variance. This process is repeated iteratively until the residual variance is minimized and the position and direction of the dipoles become stable. However, the assumptions and constraints imposed by this method are often quite strong and may not be physiologically reasonable in many cases where multiple simultaneously active sources with different locations, strengths, and orientations are present. As a result, more reasonable approaches for source localization have been developed over the last 30-40 years by multiple research groups, which is distributed source analysis.

In distributed source analysis, instead of placing a small number of dipoles, the cortical surfaces are divided into a relatively large number of voxels. This allows for the calculation of a pattern of activation values from the dipoles of each voxel, which may explain the observed voltage distribution over the scalp (Luck, 2014). This basic assumption alone cannot solve the non-uniqueness problem, due to the large number of free parameters to be estimated. To address this issue, it is necessary to incorporate neurophysiological assumptions and mathematical constraints, and to find a pattern that satisfies them. One assumption that is commonly used is that the primary sources are intracellular currents generated in the cerebral cortex, which are aligned perpendicularly to the cortical surface. This assumption is considered reasonable in most cases, given the weakened EEG signal due to the high resistance of the skull (Luck, 2014). Another solution, called *minimum norm estimation* (MNE; Hämäläinen & Ilmoniemi, 1994), is to add the constraint that the current distribution should have minimum overall source magnitudes (Hauk, 2004). During the modeling process, there are cases in which the intensity of nearby sources becomes large, which is biologically unrealistic. MNE addresses this by eliminating them through mathematical constraints.

One widely utilized method developed based on minimum norm estimation assumption is low-resolution electromagnetic tomography (LORETA; Pascual-Marqui et al., 2002) , which imposes a constraint of spatial smoothness on the solution. This approach assumes that voltage shifts gradually across the cortical surface and, as such, selects the most smooth distribution of source magnitudes. Based on these constraints, various source estimation techniques have been developed such as sLORETA (standardized LORETA; Pascual-Marqui, 2002), eLORETA (exact LORETA; Pascual-Marqui et al., 2011) and dynamic statistical parametric mapping (dSPM; Dale et al., 2000). These methods have different implementation processes depending on noise normalization or the selection of the weight matrix (Jatoi & Kamel, 2017). All of these methods have advantages and disadvantages in complexity, accuracy, and resolution, and may lead to incorrect results depending on the assumptions each method has. Therefore, rather than estimating the source only with one of these methods, it is important to cross-validate using several methods together. In this study, among several methods, MNE, dSPM, sLORETA and eLORETA, which are available in MNE-python package (Gramfort et al., 2013), are used and compared in order to obtain more accurate results.

3.5.2. Brain regions of interest

Audio-related regions: Auditory cortex

Primary and secondary auditory cortex are the sensory motor cortices involved in audio processing (Banich & Compton, 2018; Kandel et al., 2021). It is located approximately on the upper side of both temporal lobes and corresponds to Brodmann areas 41, 42 (BA 41, 42). The regions of the temporal lobe surrounding the auditory cortex (superior and middle temporal areas; BA 21, 22) are known as audio association areas and also play a role in the processing of sound (Kraemer et al., 2005). If audio words activate auditory-related regions, these regions are predicted to be involved.

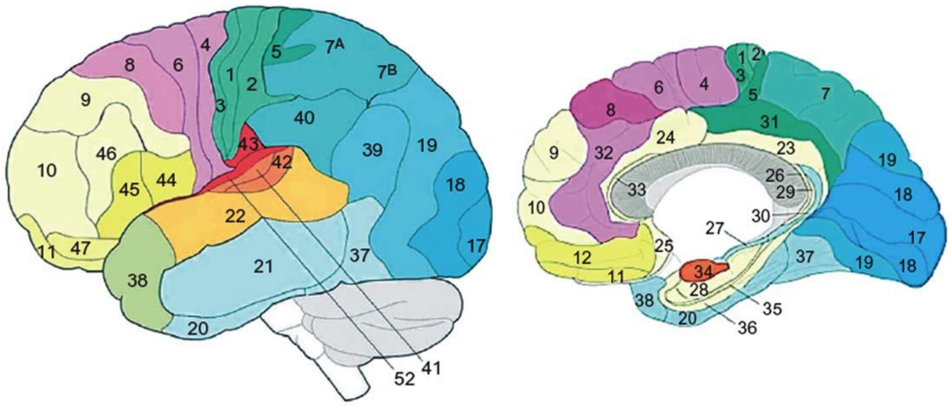


Figure 3.12. Brodmann areas. [Left] Lateral view of left hemisphere. [Right] Medial view of right hemisphere. Adapted from Gage & Baars (2018).

Interoception-related regions: insular cortex

A representative area related to interoception is the *insular cortex* (or *insula*) (Craig, 2011), which receives sensory information from the body's internal environment (Wilson-Mendenhall et al., 2019). It is concealed behind portions of the frontal, parietal, and temporal lobes inside each hemisphere's lateral sulcus (Gogolla, 2017) (Figure 5.1). Other areas known to be related to interoception are the *anterior cingulate cortex* (BA 24, 32, 33), *inferior frontal gyrus* (BA 44, 45, 47), *amygdala* (BA 25), and *somatosensory cortex* (BA 1, 2, 3), which are adjacent to the insula (García-Cordero et al., 2017; Khalsa et al., 2009). Emotion-related brain regions are *orbitofrontal cortex, anterior cingulated cortex, amygdala*, and *insula* (Phillips et al., 2003), most of which overlap with interoception-related regions. Among them, amygdala is known to be associated with fear, and insula is assumed to play a role in complex and abstract emotions (Banich & Compton, 2018).

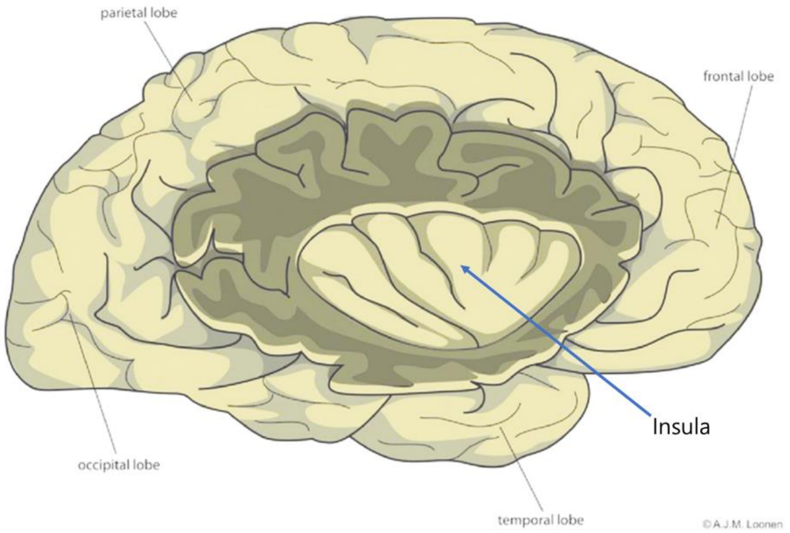


Figure 3.13. The insular cortex's location, which is hidden in the depth of the Sylvian fissure. Adapted from Loonen, Schellekens & Ivanova, (2016).

Linguistic specific regions: Broca's and Wernicke's areas

Broca's and *Wernicke's areas* are widely known classic language areas. Broca's area is located in parts of the left inferior frontal gyrus (BA 44, 45). since lesions in this area result in ungrammatical language production, it has long been associated with grammatical language production (Blank et al., 2002). Wernicke's areas were assumed to be crucial for language comprehension, as the loss of this area led to difficulties in language comprehension despite fluent articulation. It is traditionally known to be located in the posterior section of the superior temporal gyrus (STG) (BA 22) in the left hemisphere, its exact location is defined slightly vary among studies (Tremblay & Dick, 2016). Ever since the identification of these two areas, discussions of the brain areas involved in language processing were limited to these and the connections between them (Tremblay & Dick, 2016). However, studies in past several decades have shown that this classic model is not anatomically accurate and does not comprehensively account for language processing (Poeppel, 2014; Poeppel & Hickok, 2004). That language is processed solely in these kinds of linguistic areas without involvement of other areas would be evidence supporting the amodal theory.

(Amodal) semantic hub: anterior temporal lobe

Several amodal theories propose that semantic information is processed in the semantic hub regardless of modalities (Patterson, Nestor & Rogers, 2007; Visser, Jefferies, & Lambon Ralph, 2010). As for the exact location and function of these semantic hubs, there are different perspectives, but the area most frequently referred to as the semantic hub is *the anterior temporal lobe* including the temporal pole (BA 20, 21, 22, 38) (Visser et al., 2010). Damage to this area is known to cause semantic dementia. Therefore, involvement in this area would be considered as evidence supporting the amodal theory. However, it is not only amodal theory that postulates the presence of a semantic hub (Bonner & Price, 2013). Models such as the hub-and-spoke (Patterson & Ralph, 2016), which consider activation of sensory-motor areas essential in language processing, also assume a semantic hub as an area that plays a role in integrating information from sensory-motor areas or exchanging information with them (Kiefer & Pulvermüller, 2012; Patterson &

Ralph, 2016). Thus, the activation of this area with other sensory-motor regions does not contradict the prediction of embodied cognition theory.

Memory related regions: medial temporal lobe

The medial temporal lobe (BA 28, 34, 35) plays a major role in memory processing and known to be crucial for declarative memory (Raslau et al., 2015; Squire et al., 2004). It includes the hippocampus area as well as the neighboring perirhinal, entorhinal, and parahippocampal cortices. All of them are part of the hippocampus system and known to be associated with memory input/retrieval and (word) recognition (Luck et al., 2010; McCandliss et al., 2003; Tsao et al., 2018). They are also well-known as the case of H.M., who had damaged to this area, being unable to learn new, unique information (Bayley & Squire, 2005).

Chapter 4. Experiment 1

Experiment 1 aims to identify possible subgroups of linguistic concepts in terms of sensory modalities. It has purposes of classifying concepts with respect to embodiment cognition and see whether there are differences in sensorimotor perception across languages and between L1 and L2 as well as conducting a prior norming study for Experiment 2 that will investigate the sensory-specific neural responses.

4.1. Introduction

In order to enhance the theory of embodiment cognition in language comprehension, it is essential to expand the scope of research and establish a systematic classification method. Current research on embodiment cognition has provided significant evidence that language understanding is rooted in sensory-motor regions. However, the majority of this evidence was derived from the investigation of the relationship between the motor area and action verbs. This does not account for the embodiment of abstract concepts and thus, a mechanism to explain this is needed for the advancement of the theory. Furthermore, previous studies have primarily been conducted on native speakers of European languages, primarily English, with limited research on other languages, such as Asian languages such as Korean, and other types of speakers, such as L2 learners. This study aims to investigate the embodiment in L1 and L2 processing, specifically L1 Korean, L1 English, and L2 English, including both concrete and abstract concepts.

As a first step toward these purposes, the current study will construct and compare perceptual strength norms for L1 Korean, L1 English, and L2 English. Perceptual strength norms refer to a database of word profiles that describe how a word is experienced or understood in terms of the different senses. For example, a word like "computer" might be classified as a "vision" or "vision-audio" word, based on the senses through which it is typically perceived. This classification is similar to evaluating the concreteness of a word, but it is based on multiple senses rather than a concrete-abstract dichotomy. These norms can serve as appropriate stimuli when empirically testing the embodiment cognition theory. For example, the word categorization of norms can be used as stimuli when investigating how words related to vision and audio are processed differently in the brain. Additionally, the classification itself can be used to verify whether words can be captured through each sensory modality, which is an assumption of the embodiment cognition theory. Previous studies on perceptual strength norms have shown a systematic distribution rather than an arbitrary one, which is related to neurological assumptions (Reilly et al., 2020). Moreover, cognitive tasks such as lexical decision tasks based on such norms were found to be more effective than tasks with simple binary criteria such as concreteness and imageability (Connell & Lynott, 2016). The current study intends to construct perceptual strength norms for these two purposes. The constructed norms will be used as stimuli for Experiment 2 and will also serve as a means for comparison of embodiment aspects between languages.

Sensorimotor norms have been established in a variety of languages, including English (Lynott et al., 2020), Russian (Miklashevsky, 2018), Dutch (Speed & Majid, 2017; Speed & Brybaert, 2021), Italian (Vergallito et al., 2020), French (Chedid et al., 2019; Miceli et al., 2021), Serbian (Đurđević et al., 2016), and Mandarin (Chen et al., 2019). However, no such investigation in Korean or for L2 learners was conducted. According to the embodiment theory, the experience of language or the environment surrounding language has an important influence on language embodiment. That is, the same concept might be differently embodied for Korean and English speakers, or L1 and L2 speakers. From this point of view, the sensory experience of a word, or the perceptual strength norm, may vary from language to language or depending on whether it is L1 or L2. Therefore, perceptual strength norms for Korean and L2 English learners should be preceded for behavioral and neurophysiological experiments. Moreover, most studies have focused on accumulating perceptual strength norms in each language, not attempting to compare them across languages. This study will not only collect Korean and L2 English data but also compare them with L1 English data to examine the differences across languages.

There are largely two methods to measure the sensorimotor perception of concepts: feature listing (or property generation task) and rating methods (Conca, Borsa, Cappa, & Catricalà, 2021). In feature listing tasks, participants are asked to generate a list of features or a definition to explain a certain notion. Then based on this generated list, concepts are classified. This method has an advantage in revealing many properties of words that are otherwise easily overlooked, by allowing participants to freely associate with words. However, this process of free association also acts as a disadvantage, when the task includes abstract concepts, because the heterogeneous nature of abstract concepts is highly likely to present excessively diverse results. The experimenter's subjectivity should be involved in filtering those many properties, which might not be suitable for systematic use.

Moreover, filtering stage is time and effort consuming, thus difficult to apply to a large number of words. On the other hand, in rating methods, participants evaluate on a Likert scale how relevant a given concept is with several semantic dimensions pre-selected by the experimenter. Compared to feature listings, rating methods are less time demanding, so they allowed researchers to gather data for a larger number of stimuli and from a number of individuals. However, the dimensions for ratings are often chosen with arbitrary but no empirical or theoretical reason, which might lead to potentially biased results.

The present study will employ dimension rating methods, since it is more appropriate to collect data from diverse group of individuals. This study requires at least three sensorimotor norms—L1 Korean, L1 English and L2 English. So, since three classification tasks are required in this study, the rating method advantageous for processing large amounts of data is proper. Moreover, the present study will compare the classification results across each language group. A rating method is suitable for this purpose in that it quantifies the semantic dimension based on a preset criterion, which facilitates comparison between languages.

In order to ensure unbiased and organized outcomes, it is crucial to select an appropriate criterion for evaluating the dimension when utilizing the rating method. This study will base the criterion on the sensorimotor norms in Lynott et al. (2020), which assessed the sensorimotor strength of approximately 40,000 English words from 3,500 participants. It employed six perceptual modalities (touch, hearing, smell, taste, vision, and interoception) and five action effectors (mouth/throat, hand/arm, foot/leg, head excluding mouth/throat, and torso) as the semantic dimensions. This classification system has two key benefits. Firstly, it is based on some extent on neuroanatomical evidence. The primary somatosensory cortex, which accepts the information from peripheral receptors and translates into meaningful representation in the brain (Raju & Tadi, 2020), consists of distinct cortical regions corresponding to each sensory modality. For example, primary auditory cortex is seen on the top boundary of the temporal lobe while the primary visual cortex is found in the occipital lobe’s extreme posterior tip (Marieb & Hoehn, 2007). Thus, the classification system by each sense and body-effector reflects the neuroanatomical background. Secondly, it includes interoception as one of the modalities in addition to the commonly used five senses. Interoception is an important sensory modality particularly for researching the embodiment of abstract concepts since it is likely to be related with them (Connell et al., 2018; Critchley & Harrison, 2013). This study aims to classify vocabulary based on systematic categorization with a neurological foundation and apply this criterion identically to both concrete and abstract concepts. Because the advantages of Lynott et al. (2020)'s semantic dimensions align with this goal, this study will adopt them as a criterion for the perceptual strength norms.

4.2. Method

4.2.1. Materials

A total of 1000 English words were selected from about 40,000 words used in The Lancaster Sensorimotor Norms (Lynott et al., 2020) considering familiarity, word frequency, max perceptual strength, exclusivity, and translatability. On the basis of these criteria, words were chosen that are well understood by L2 English speakers, are straightforward to translate into Korean, and may serve as the experimental material for experiment 2. Frequency and familiarity are factors for selecting words that L2 participants can understand relatively easily. Max strength refers to the score of the sense with the highest one among the six senses (visual, auditory, haptic, gustatory, olfactory, and interoceptive) in Lynott et al. (2020). Exclusivity is the extent to which a concept is experienced through a single sense, calculated by dividing the sum of all sensory scores by the max strength. Since this study works as perceptual norms for Experiment 2 and future neurolinguistic studies, the items more likely to reveal the difference in neural activity are selected. Words with similar scores in several senses (low exclusivity) or words with low perceptual strength (low max strength) may not be appropriate to elicit different neural activities. First, two-word items and words with less than 90% familiarity by native speakers were removed. Then, words with frequency of less than 3.5 on the Zipf scale[2] were excluded. Word frequency was measured using the Python package WordFreq (Speer et al., 2018). This package uses the Exquisite Corpus, which encompasses Wikipedia, books, web and other sources, so it is likely to reflect a more realistic frequency for present-day speakers. After this filtering process, about 9,000 words were left. Among these words, those with below-average max strength and exclusivity within each sense were excluded in order to select words with higher max strength and exclusivity were selected. Lastly, words that cannot be easily translated into Korean were eliminated, such as polysemy, prepositions, pronouns, and onomatopoeia. Selected words were composed of several parts of speech (nouns, verbs, adjectives) and various concepts (e.g., food,

[2] Zipf frequency is the base-10 logarithm of the number of times it appears per billion words. For example, a word with Zipf value 6 appears once per thousand words and a word with Zipf value 3 appears once per million words (Van Heuven, Mandera, Keuleers, & Brysbaert, 2014).

animals, emotions, sports, occupations, colors, etc.). They were randomly divided into 10 sets of 100 words. This was to break it down into sets suitable for evaluation during single experiment session, since it is difficult for one participant to evaluate all the words at once.

A Korean set consisted of Korean translation equivalents of the English set. Translations into Korean were created based on the experience of language users rather than dictionary definitions. This was to narrow the gap between the meaning of a word commonly recognized by L2 users and its dictionary definition, as well as to meet the purpose of this experiment measuring the sensory experience of a word. For the translation process, 11 participants were recruited as translators. Since they were also L2 learners who had similar levels of English proficiency, academic background, and age as the participants in the main experiment, they were expected to have similar sensory experiences for words. Each of them translated 1000 English words into Korean. They were instructed to translate each English word into the Korean word that sprang to mind, that is, to use their experience of words. If a word has multiple meanings with several possible translation candidates, they were forced to choose the first one that they recalled. Translations into loanwords were permitted. As they are commonly accepted as part of Korean vocabulary, there is no need to exclude them. Moreover, they can serve as a great testbed for examining differences between languages due to their close correspondence with the original vocabulary. Nevertheless, when their first translation was a loanword, they were asked to come up with another possible translation that was not a loanword and to select the one that was more common in the context of everyday life between them. This was to rule out the cases that the translated words were more likely to be recognized as a foreign language rather than Korean or simple transliteration of English words. In addition, they were encouraged to take a part-of-speech into account during translation. If a word is represented as more than one part-of-speech, they had to choose the one that they thought was more frequently used. They were not allowed to look up a dictionary or other reference materials. When they did not know the meaning of a word, they were instructed to skip it over. After collecting the translation candidates, the translation agreed upon by the most translators was selected as the representative translation. If more than one translation were mostly agreed, the one closer to the first definition in the English-Korean dictionary was chosen by the experimenter's judgment. In the case that different English words were translated into the same Korean word, (*e.g.* angry / pissed “화난”), one of the translations was replaced with the other translation candidate or a Korean synonym, matching the frequency of Korean and English (angry as “화난” and pissed as “열받은”). Meanwhile, if the Korean translation has ambiguity, cues that can solve it are additionally presented (eyes as “눈 (신체의 일부)”). Although carefully designed, some of the Korean translations do not obviously correspond with English words. This limitation might make it difficult to have accurate statistical comparisons between languages. However, about 60% of the translations were agreed among more than half of the translators, and the translations with less than 2 translators agreeing upon account for only about 10% of the words. Therefore, although it might not be an accurate statistical comparison, it seems to be a reasonable level to confirm the difference in the overall distribution between languages.

4.2.2. Procedure

The surveys were created using PCIbex (Zehr & Schwarz, 2018) and conducted online. It proceeded as follows. First, participants read about the outline of the experiment and started it if they agreed to participate in it. Then, they learned the procedure, and after five practice trials, they performed the main task. The experiment consisted of 100 words, which were randomly ordered in each participation. It took an average of 25 minutes for the completion. The tasks for Korean participants including L2 English survey were presented in Korean to help participants understand accurately.

During the task, participants were asked to evaluate the sensory experience and familiarity on English or Korean words. Each word first appeared on the screen, followed by rating scales located below the word. Above scales were the labels of senses evaluated. Participants rated how strongly they used six senses in experiencing a particular word. The six senses were visual, auditory, haptic, gustatory, olfactory, and interoceptive. Interoceptive sense is the one felt through the receptor organs inside the body, such as body temperature, hunger, thirst, digestion, and heartbeat, which was explained in the instruction and during the task (Figure 4.1). Through the Instruction, it was explained that there is no right or wrong answer, so evaluation should be based on one’s own intuition. A 6-point Likert scale was used ranging from 0 (not experiencing the sense at all) to 5 (experiencing the sense strongly). After a perceptual strength evaluation, familiarity was rated. Participants assessed how familiar they were with the word, or how well they know its meaning, on a 4-point Likert scale (not at all 0 - very well 3).

Figure 4.1. An example of the evaluation task. Participants selected the number either by clicking or touching. They were able to expand the explanation on interoceptive sense whenever they wanted to read it.

The five practice trials presented before the main task were not only a means to assist in understanding of the task, but also functioned as calibrator words. They played a role in allowing participants to employ various sensory dimensions over the entire range through words with relatively clear sensory indices. The same five calibrators with those in Lynott et al. (2020) were used, which were *account* (Low strength across all modalities), *breath* (medium strength across multiple modalities), *echo* (high strength in a single modality), *hungry* (uneven strength across modalities) and *liquid* (high strength across multiple modalities).

4.2.3. Participants

For L2 English dataset, a total of 55 unique participants[3] (male: 21, female: 34) were recruited for the experiment. They completed on average 3.51 lists each. They were Korean learners of L2 English, who were born in Korea. Their average age was 26.78 years old. Their English proficiency level was advanced according to TEPS[4], and the average TEPS score was 417. Participants voluntarily participated in the experiment, recruited from the university website and were paid 5,000 won per list as a reward for participation. They were allowed to participate in more than one word list from a total of 10 lists, but not in the same list twice. Each word list was rated by 19.3 participants on average, and all words were rated by at least 18 participants. 48 participants (male: 19, female: 29) completed the L1 Korean survey, and each evaluated 4.40 lists on average. Their average age was 26.51 years and their average TEPS score was 427.5, which were not significantly different from the mean age and TEPS score of L2 English participants ($p > 0.1$). Each word list was rated by an average of 21 participants, each rated by at least 19 participants. Others were the same as for L2 survey. For L1 English survey, 173 unique participants (male: 93, female: 80) were recruited from Amazon Mechanical Turk. Participation was restricted to only those who speak English as their first language and reside in an English-speaking country. They were paid $4.2 for their participation, which was roughly equivalent to 5,000 won. They performed an average of 1.04 list and their average age was 40.38 years. Compared to the other two surveys, the L1 English survey was distinguished in that the participants were older and had different academic backgrounds (All of

---

[3] The number of participants does not include those who seemed to randomly pick the options.

[4] Test of English Proficiency developed by Seoul National University.

participants of L2 English and L1 Korean studies were undergraduate or graduate students), and a larger number of participants participated. Each word was rated by an average of 19.1 participants and by at least 18.

4.2.4. Analysis

The analysis included both quantitative and qualitative analysis. Quantitative analysis aims to identify similarities and differences between languages or modalities by exploring the perceptual strengths of words in various ways. In the following *results* section, the findings of interrater reliability, perceptual strength, dominant modality, exclusivity, relationship between modalities and relationship between modalities and concreteness will be presented and briefly discussed. For quantitative analysis, the collected data was trimmed as follows.

For L2 English data, 29 words were excluded from the analysis, due to low average familiarity (below 2). So, a total of 971 words were analyzed. The individual rated items with low familiarity (below 2) were also removed from the analysis, which accounted for 3.85% of the data. In L1 Korean, 11 words were removed, and 3.41% of items were excluded from the analysis. In L1 English no word had a familiarity of less than 2 and 1.12% of the data were removed from the analysis.

A comparative analysis of individual words was also conducted to analyze the differences between languages that were not likely to be revealed by quantitative analysis alone. Several intriguing cases with implications for linguistic differences or embodiment processing were selected through visual inspection on the radar charts of perceptual strengths (Figure 4.2), considering various factors such as polysemy, usage and loanwords. The findings will be described in the *individual cases* and *individual cases for loanwords* section.

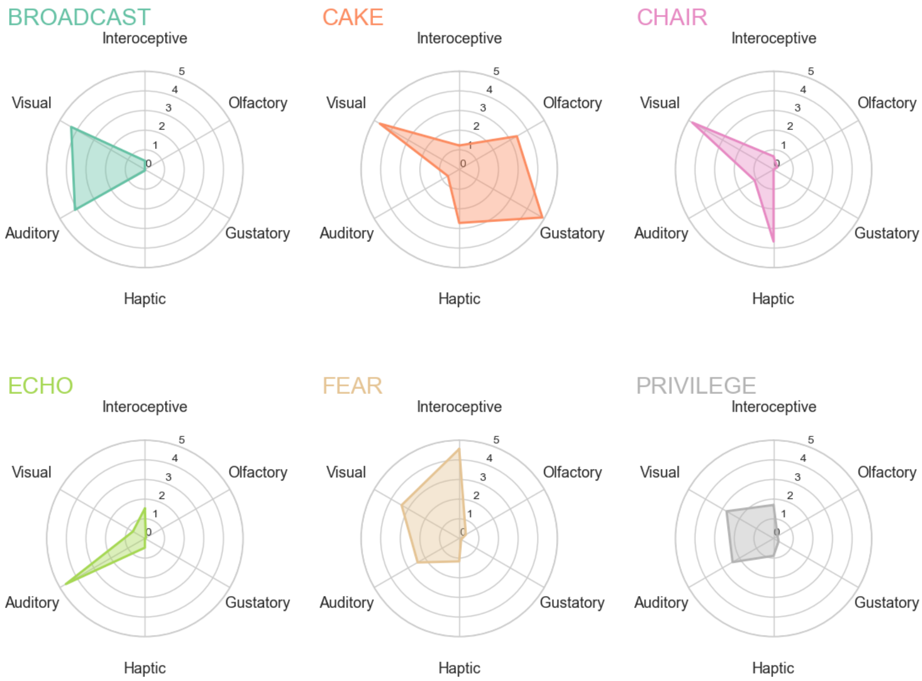


Figure 4.2. Radar charts for individual concepts (examples from L1 English data). It displays the mean perceptual strengths on each of the 6 dimensions. Some words such as *echo* are exclusively related with one modality, whereas other words such as *broadcast*, *cake*, and *chair* are experienced through two or more modalities. *Fear* is one of the examples rated high in interoception, while *privilege* is the case where all modalities were evaluated low.

4.3. Results

4.3.1. Interrater reliability

Interrater reliability for perceptual strengths was calculated using Cronbach's alpha, which is a measure of internal consistency. Cronbach's alpha was calculated per item list and then averaged within each language, using a python statistical package, Pingouin (Vallat, 2018). All language groups showed excellent interrater reliability for all modalities ($\alpha > 0.9$) (Table 4.1-a). Each Cronbach's alpha level in this study was comparable to that of Lynott et al. (2020). This suggested that perceptual strengths evaluated by participants were not arbitrary but internally consistent, reflecting what they evaluated.

4.3.2. Perceptual strengths

Summary statistics (mean and standard deviations) per modality are shown in Table 4.1-b and Figure 4.3 and selected graphical examples in Figure 4.2. There seemed to be no noticeable difference in the overall mean and distribution of perceptual strengths in L1 Korean, L1 English and L2 English. In all languages, visual dimension was rated the highest, while olfactory and gustatory were low. This distribution is almost similar to that from other perceptual strengths norm in previous studies. For example, in of Lynott et al. (2020), visual rating was also the highest, auditory second highest, and gustatory and olfactory the lowest (Auditory: 1.51, Gustatory: 0.32, Haptic: 1.07, Interoceptive: 1.03, Olfactory: 0.39, Visual: 2.90). Although the perceptual strengths of this study were overall larger than those of Lynott et al. (2020), particularly visual strength was higher by an average of 1 point, this was expected because the material in this study were the items with high max strength selected from Lynott et al. (2020). Nevertheless, smell, taste, and interoception were almost similar to the overall mean of 40,000 English words in Lynott et al. (2020).

When perceptual strengths were statistically compared across languages[5] there were no significant differences between languages in vision, audio, and touch (*p>.05*), while significant differences in interoception (t=*8.31, p<0.001*), smell (*t=9.64, p<0.001*), and taste (*t=6.3572, p<0.01*). Within interoception, L1 Korean and L2 English showed significant difference (*t=3.650, p<0.001*), while in taste, perceptual strengths between L1 English and L2 English were significantly different (*t=3.037, p<0.01*). On the other hand, in olfactory modality, three language groups were different from each other (L1 English vs L2 English: *t=2.079, p<0.05*, L1 English vs L1 Korean: *t=2.826, p<0.01*, L1 Korean vs L2 English: *t=3.955, p<0.001*).

		Auditory	**Gustatory**	**Haptic**	**Interoceptive**	**Olfactory**	**Visual**
(a)		**Interrater Reliability (Cronbach's alpha)**					
Korean	L1	0.98	0.95	0.96	0.98	0.95	0.98
English	L1	0.98	0.92	0.96	0.97	0.93	0.96
	L2	0.98	0.98	0.98	0.98	0.98	0.98
(b)		**Perceptual Strength Ratings**					
Korean	L1	1.74 (1.77)	0.29 (0.90)	1.26 (1.61)	0.94 (1.49)	0.37 (0.92)	3.91 (1.33)
English	L1	2.06 (2.00)	0.28 (0.95)	1.57 (1.88)	1.06 (1.66)	0.33 (0.95)	3.93 (1.59)
	L2	1.69 (1.70)	0.33 (0.88)	1.29 (1.54)	1.16 (1.51)	0.40 (0.89)	3.70 (1.35)
(c)		**Correlations between languages**					
L1E *vs.* L2E		0.81***	0.91***	0.82***	0.83***	0.83***	0.81***
L1E *vs.* L1K		0.74***	0.87***	0.73***	0.80***	0.80***	0.75***
L2E *vs.* L1K		0.88***	0.90***	0.85***	0.84***	0.84***	0.84***

Table 4.1. Statistics for perceptual strengths. (a) Mean Cronbach's alpha for each modality within each language group. Larger than 0.9 generally means excellent internal consistency. (b) Mean perceptual strength ratings (0–5) and standard deviations (SD). (c) The correlation scores across languages. Larger than 0.7 is generally considered strong. All coefficients are significant (*p<0.001*).

5 It was calculated with mixed effect linear regression model using R statistical packages, lme4 (Bates et al., 2015) and lmerTest (Kunzetsova et al., 2017).

The correlations between languages for each modality were measured to see whether language groups had similar distribution. The correlations were computed by using as variables the mean perceptual strengths for each modality per word (average of perceptual strengths rated by each participant). As a result, there was overall a strong correlation between languages in all modalities (Table 4.1-c). In general, the correlation coefficients between L1 English and L1 Korean were lower than L2 English and L1 Korean or L1 English and L2 English.

In addition, Euclidean distances between languages were calculated to measure similarities between language groups. The Euclidean distance is the way to determine the distance between two points, often used as a measure of similarity. The distances between languages were computed in a pairwise manner, resulting in a vector of the distances for the mean perceptual strengths of each word between two language groups. For instance, the mean perceptual strengths of the word, *addiction*, were [auditory: 1.5, Gustatory: 1.4, Haptic: 1.11, Interoceptive: 3.22, Olfactory: 0.89, Visual: 3] in L1 English, and [auditory: 1.26, Gustatory: 0.74, Haptic: 1.89, Interoceptive: 2.95, Olfactory: 1.58, Visual: 2.58] in L2 English. The distance between these two points is calculated as 0.93. These word distances can be obtained for all words in two language groups, which make up a vector of the word distances. The mean of the word distances between languages were relatively small for all comparisons (L1 English vs L2 English: 1.22, L1 English vs L1 Korean: 1.38, L1 Korean vs L2 English: 0.97), compared to the distances between each language and randomly created perceptual strengths[6], which were about 5 on

[6] It was created by drawing samples of the same number of words in each language from a uniform distribution. Each sample consisted of six values, selected within the range of 0.0

average. This indicated that there were similarities between languages. L1 Korean and L1 English showed the lowest similarity among them, which was consistent with the findings that they had the weakest correlations. On the other hand, L1 Korean and L2 English had the highest similarity. If interpreted simply, L2 English was more similar to L1 Korean than L1 English.

In summary, the distributions of perceptual strengths in each language were similar to those in previous studies—higher visual rate and lower olfactory and gustatory. Also, although there were some differences in each rating, considering the correlation and Euclidean distance, the overall distribution across all language groups seems to be similar.

4.3.3. Dominant modalities

The dominant modality for a word is the one with the highest rating among six sensory modalities (Table 4.2). When two modalities tied the highest rating, both were treated as dominant modalities. For example, as *finger* was rated as 4.10 both in haptic and visual sense by L2 English participants, it was labeled as Haptic-Visual dominant words and counted respectively as a haptic-dominant word and a visual-dominant word.

The distributions of dominant modality were also quite similar across languages. For all language groups, visual sense was the most dominant sensory dimension, while the proportion of olfactory, haptic and gustatory was relatively low. For gustatory, even though the average of perceptual strength was lower than

to 5.0 to simulate the mean perceptual strengths of six modality. For example, L1 English group has 1000 words, so the random sample dataset is constructed with 1000 six random values.

that of haptic, the proportion of dominant modality was larger.

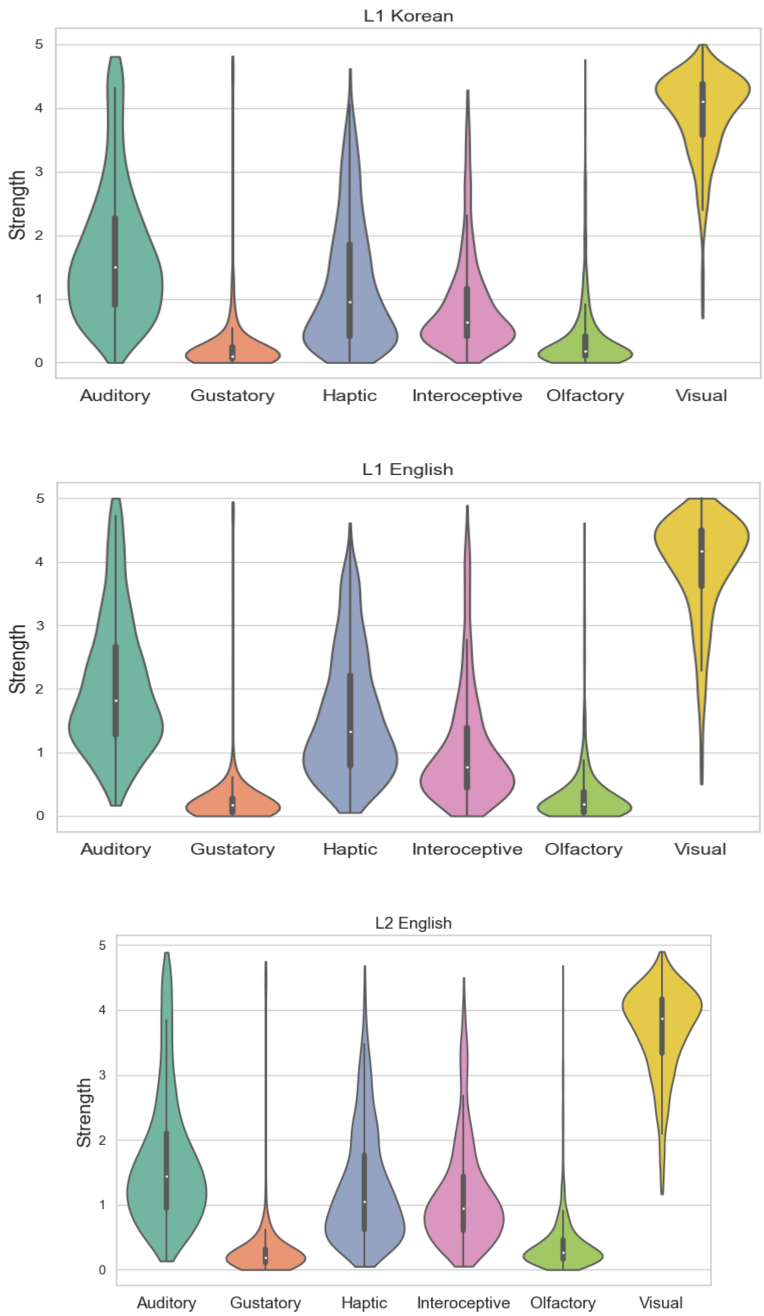


Figure 4.3. Violin plots showing the distribution of perceptual strength ratings for each sensory modalities (Top: L1 Korean, Middle: L1 English, Bottom: L2 English). Black lines and boxes inside indicating boxplots. The three groups have similarity in shape and distributions.

While the three language groups showed an overall similar distribution, L1 Korean appeared to be distinguished in that it was more visually dominant and less dominant in the haptic and interoception, which was consistent with lower perceptual strength means in interoceptive and haptic sense in L1 Korean, although the difference was not statistically significant. This dominance tendency of L1 Korean, strong in vision while weak in interoception and haptic, was also revealed from the findings that words rated as haptic and interoceptive words in L1 English and L2 English were mainly evaluated as visual words in L1. For example, 33 out of 60 interoceptive dominance words in L1 English had visual dominance in L1 Korean except for one word (32 out of 59 interoception words in L2 English → 31 vision and 1 haptic in L1 Korean). In contrast, there were only few cases that interoception-dominant words in L1 Korean had other modality dominance in L1 English (5 cases) and L2 English (2 cases). However, those with different modality in L1 Korean were not experienced in a completely different way compared to L1 English and L2 English. Although several interoceptive and haptic words in L1 English and L2 English were rated as vision-dominant in L1 Korean, the interoception and haptic strengths of those words in L1 Korean were considerably higher than the average of the interoception and haptic while the vision strengths of those words in L1 English and L2 English were greater than the average. For example, interoception strengths of those words were 2.87 on average, which was lower than 3.69 in L1 English but much higher than the overall average in interoception of 0.94. Also, the average visual strength of these words was 3.44, which was higher than the 2.73 of L1 English but lower than the overall average of strengths in vision, 3.91. In short, the words with haptic or interoceptive dominance in L1 English and L2 English were often visually dominant in L1 Korean, but their profile of perceptual strengths was not completely different from the other language groups.

L1 English and L1 Korean had different modality in 51 cases (L1E≠L1K), L1 English and L2 English in 48 cases (L1E≠L2E), and L1 Korean and L2 English in 66 cases (L1K ≠L2E). Most of the words dominated by other modalities in one language were found to be visually dominant in the other languages (e.g. Auditory in L1 English → Visual in L2 English) or vision tied for the highest rating (e.g. Interoceptive in L1 English → Interoceptive & Visual in L2 English). Also, the second dominant modality for most of these words was consistent with the dominant modality of the other languages (e.g. Haptic dominant in L1 English → Vision as 1st dominant / Haptic as 2nd dominant in L1 Korean). In this respect, all language groups seemed to have a similar distribution in dominant modalities.

		Aud.	**Gus.**	**Haptic**	**Intero.**	**Olf.**	**Visual**
Korean	**L1**	73 (7.97%)	15 (1.64%)	6 (0.66%)	32 (3.49%)	1 (0.11%)	862 (87.16%)
English	**L1**	98 (9.73%)	18 (1.79%)	11 (1.09%)	60 (5.96%)	3 (0.30%)	817 (81.13%)
	L2	81 (8.28%)	17 (1.74%)	10 (1.02%)	59 (6.03%)	2 (0.20%)	809 (82.72%)

Table 4.2. Numbers and percentage of dominant modalities within L1 Korean, L1 English and L2 English. *Aud.=auditory, Gus.=gustatory, Intero.=interoceptive, Olf.=olfactory*

The cases that L1 English and L2 English had the same dominant modality but L1 Korean had different dominant were 47 (L1E=L2E≠L1K), those with the same dominance in L1 Korean and L2 English but different in L1 English (L1E≠L2E=L1K) were 27 different cases, those with the same in L1 English and

L1 Korean but not in L2 English were 16 (L1E=L1K≠L2E). The most frequent case was the one that words were evaluated as different modality only in L1 Korean (L1E=L2E≠L1K). This might be because Korean and English did not have a complete one-to-one correspondence for some words, which inevitably made L1 Korean participants have different sensory experiences. For instance, a polysemy such as "hard" has several meanings in English, such as solid (mostly haptic) or difficult (possibly interoceptive) that can be experienced with different senses. However, as it was translated into one of the meanings in Korean ("어려운 difficult"), it was not easy to be evaluated as haptic words. On the other hand, it was less common for L1 English and L1 Korean to provide the same evaluation but L2 English to give a different one (L1E=L1K≠L2E). This might be because L1 English and L1 Korean have the fewest cultural or linguistic connections between them while L2 English must have more links either with L1 Korean or L1 English. Most of the cases in which only one language group showed different modality were related to vision and the second most dominant modality was mostly in agreement with the dominant modality of the other groups.

There were only two words where all language groups had a different dominant modality from each other: thermal ("열의") and calm ("차분한"). This difference seemed to be due to the failure to resolve the ambiguity of Korean[7] and the lack of one-to-one correspondence between Korean and English.

[7] A Korean translation for thermal, "열의" has two different meanings—thermal and enthusiasm— which was not considered in the material design phase.

4.3.4. Exclusivity

Exclusivity scores per word are calculated by dividing the max perceptual strength with the sum of all strengths. It ranges from 0 % (experienced equally in all senses) to 100% (solely through a single sense). A mean exclusivity score for L1 English was 46.37%, for L2 English 46.71% and for L1 Korean 49.84% (Table 4.3). According to the average exclusivity, L1 Korean was less multi-dimensional than the other two groups. Their average exclusivity was generally higher than 43.5% in Lynott et al. (2020), which was expected considering that materials with high exclusivity were selected from it. The correlation coefficient of exclusivity between L1 English and L2 English was 0.67 ($p<0.001$), between L1 Korean and L2 English was 0.75 ($p<0.001$), between L1 English and L1 Korean was 0.59 ($p<0.001$). The correlations were generally strong, and the lowest between L1 English and L1 Korean, similar to those in perceptual strengths.

In L1 English, *rainbow* was the least muti-modality word, with 82.46% exclusivity, while *drinking* was the least exclusive one, with exclusivity of 20.7%. Five most exclusive words were *rainbow*, *white*, *online*, *mood*, and *download*, which were all visual words. Five least exclusive words were *drinking*, *salad*, *lemon*, *apple* and *delicious*, all gustatory words. Because they were often accompanied by vision, smell, or interoception, the exclusivity score was low. In L2 English, *brown* was the word with the highest exclusivity score (80.73%), and *feel* was the word with the lowest exclusivity score (20.78%). Five highest words were all visual words (*brown, white, gray, moon, and yellow*), while five lowest words were *feel, drinking, healing, pizza,* and *honey*, most of which were related with gustatory sense but also included haptic (*feel*), visual (*drinking*), and interoceptive (*healing*) words. In L1 Korean, five most exclusive words were *shadow* (“그림자” 88.69%), *star* (“별”), *moon* (“달”), *rainbow* (“무지개”), and *graph* (“그래프”), whereas least exclusive ones were *feel* (“느끼다” 20.72%), *drinking* (“마시는”), *happy* (“행복한”), *pizza* (“피자”), and *apple* (“사과”). The most exclusive words were all visual dominant words, while the least ones were interoceptive (*feel*), gustatory (*drinking*, *pizza*), and visual (*happy*, *apple*) words with high gustatory strengths.

	Exclusivity	**10 most exclusive words (dominant modalities)**	**10 least exclusive words (dominant modalities)**
L1 Korean	49.84% (*sd=11.08*)	shadow, moon, star, rainbow, graph, satellite, visual, cloud, cloud, lane (visual)	feel, hunger, pulse (interoceptive), drinking, pizza, cake, honey, butter (gustatory), happy, apple (visual)
L1 English	46.37% (*sd=8.47*)	rainbow, white, online, moon, download, percentage, galaxy, email, mars, shadow (visual)	drinking, salad, lemon, apple, delicious, cake, honey, olive, juice, pizza (gustatory)
L2 English	46.71% (*sd=9.10*)	brown, white, gray, moon, yellow, green, horizon, logo, black, diameter (visual)	feel (haptic), drinking (visual), healing (interoceptive), pizza, honey, apple, salad, addiction, cherry, rice (gustatory)

Table 4.3. 10 most exclusive words and 10 least exclusive words for L1 Korean, L1 English, and L2 English. Listed first are the most exclusive or least exclusive words, but the rest may not be in order depending on dominant modalities.

The most exclusive words were visually dominant in all three language groups. The five words with high scores in each group overlapped each other, and even if not, they all had high exclusive scores. The least exclusive words also had cross-linguistic overlap and were similar in that they were in most cases related to gustatory sense. However, L1 Korean and L2 English included *feel*, *healing*, and *happy*, which had less relevance with gustatory sense and more with interoception.

In L1 English they were also rated low exclusivity (34.22%, 34.92%, 37.05% each) but not as low as those in L1 Korean and L2 English all below 30%.

4.3.5. Relationship between modalities

The correlations between sensory modalities were calculated (Table 4.4). Similar to previous studies (Connell & Lynott, 2012; Vergallito et al., 2020), gustatory and olfactory strength were highly correlated, while visual and haptic strength had modest correlation. This was consistent with our intuition that one can experience both taste and smell in the intake of food and touchable objects are often visible. On the other hand, auditory sensory showed a moderate negative correlation with vision and touch, while interoception was negatively correlated to vision. This was also compatible with our observation that audible objects cannot be seen or touched, nor can we see sensations inside the body.

In addition to this, there were other significant correlations, but they were inconsistent across language groups or the size of association was small. There was a weak correlation between interoception and gustatory, between haptic and olfactory and between haptic and gustatory for all language groups. Since all of these senses had low strengths on average (many of words were rated as near-zero for these senses, particularly for olfactory and gustatory), these weak correlations may be driven by the relationship between only several words. However, since they were consistent with previous studies (Connell & Lynott, 2012; Vergallito et al., 2020), it may be helpful to explore the data in detail. For the correlation between interoceptive and gustatory, the explanatory interpretation may be added that what can be tasted was also related to the senses of the digestion, a type of interoception, or something tasty brought emotional experience.

		Aud.	**Gus.**	**Haptic**	**Intero.**	**Olf.**	**Visual**
Aud.	L1K		-0.06	-0.29***	0.12***	-0.02	-0.32***
	L1E		-0.17***	-0.41***	0.07	-0.19***	-0.36***
	L2E		-0.08	-0.32***	0.11***	-0.07	-0.31***
Gus.	L1K	-0.06		0.17***	0.15***	0.73***	-0.05
	L1E	-0.17***		0.1**	0.11***	0.76***	-0.08**
	L2E	-0.08		0.17***	0.15***	0.75***	-0.04
Haptic	L1K	-0.29***	0.17***		0.04	0.22***	0.37***
	L1E	-0.41***	0.1**		-0.11***	0.17***	0.35***
	L2E	-0.32***	0.17***		-0.04	0.19***	0.34***
Intero.	L1K	0.12***	0.15***	0.04		0.1**	-0.34***
	L1E	0.07	0.11***	-0.11***		0.06	-0.51***
	L2E	0.11***	0.15***	-0.04		0.15***	-0.42***
Olf.	L1K	-0.02	0.73***	0.22***	0.1**		0.06
	L1E	-0.19***	0.76***	0.17***	0.06		0.03
	L2E	-0.07	0.75***	0.19***	0.15***		0.03
Visual	L1K	-0.32***	-0.05	0.37***	-0.34***	0.06	
	L1E	-0.36***	-0.08**	0.35***	-0.51***	0.03	
	L2E	-0.31***	-0.04	0.34***	-0.42***	0.03	

Table 4.4. The correlation matrix between modalities. Orange shade indicates negative correlation and blue positive, with darker color being stronger correlation. **:p<0.05, **: p<0.01, ***: p<0.001, Aud.=auditory, Gus.=gustatory, Intero.=interoceptive, Olf.=olfactory*

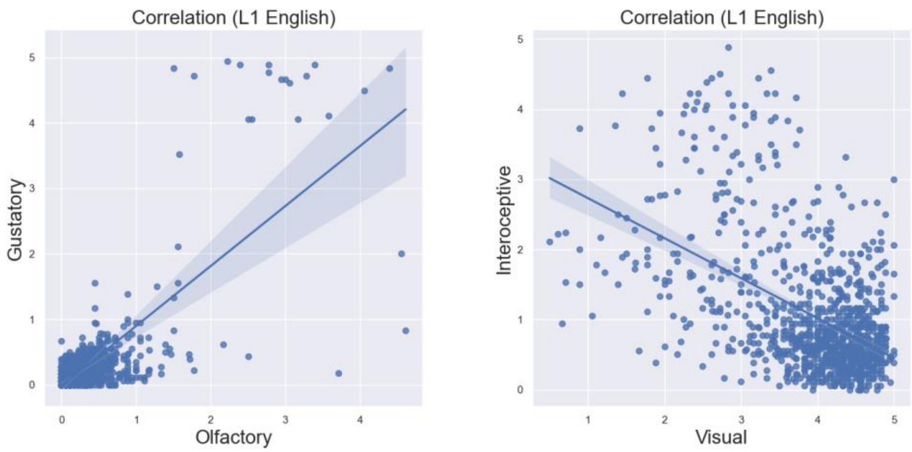


Figure 4.4. Correlation between two modalities. [Left] The scatter plot with regression fit line between Gustatory and Olfactory in L1 English. [Right] Visual and Interoceptive in L1 English.

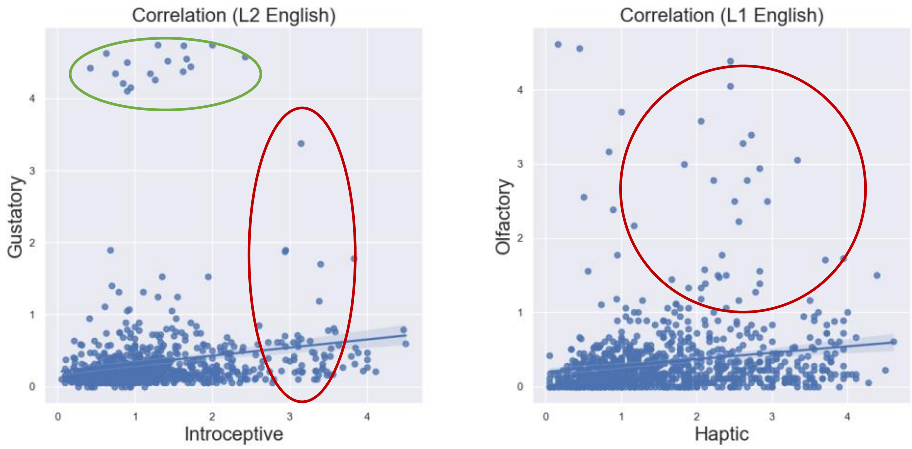


Figure 4.5. Weak correlation between two modalities. [Left] Interoceptive and Gustatory in L2 English. The words that might contribute to the correlation are circled in red. [Right] Olfactory and Gustatory in L1 English. Some words that might contribute to the correlation are circled in red.

However, it was difficult to find evidence to support this explanation from the data. As shown in Figure 4.5 (Left), a cohesive pattern was not revealed, but rather, high gustatory words seemed to form separate clusters (green circle). In general, the weak correlation appeared to be caused by overall low strengths in the two senses and the relationship between certain words. Such words were *drinking*, *addiction*, *hunger*, *feel*, *happy*, and *healing*, the ones with high interoception strengths and related to eating behavior (red circle). Olfactory and haptic seemed to be also weakly correlated due to overall low strengths and the relationships of several words but have a more cohesive pattern compared to that of interoception and gustatory (Figure 4.5. Right). The words included in a red circle area were *sugar, olive, pizza, apple, butter,* and *shore*. These were the words that could be both touched and smelled although neither haptic nor olfactory was the primary sense for experiencing them—most of them were gustatory words. A similar explanation can be provided for the weak correlation between haptic and gustatory.

A few words relevant both with gustatory and haptic sense, which in most cases overlapped with above example of haptic and olfactory, induced their correlation.

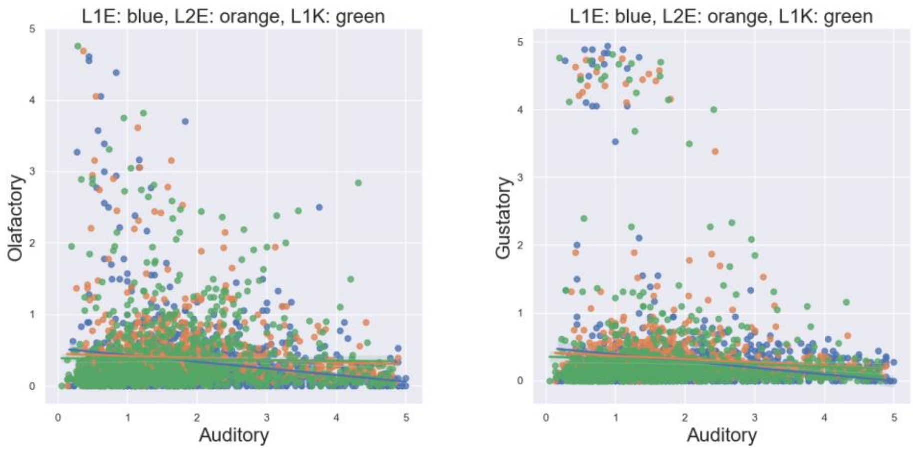


Figure 4.6. Cross-linguistic comparison of correlation between two modalities [Left] The scatter plot with regression fit line between Auditory and Olfactory in L1 English (blue), L2 English (orange), and L1 Korean (green). [Right] Auditory and Gustatory.

Auditory sense had a weak negative correlation with gustatory and olfactory only in L1 English group. The negative correlations between them were also found from English speakers (Connell & Lynott, 2012) and Italian speakers (Vergallito et al., 2020), but not from Korean speakers and L2 English learners in this study. Although something that can be tasted or smelled is unlikely to be experienced through auditory senses so that it is plausible to find negative correlation between them, the absence of a negative correlation between the sense of taste or smell and the sense of hearing is also not unexpected due to the possibility of no correlation, positive or negative, existing between the two senses at all. Moreover, since gustatory and olfactory were evaluated as near-zero in many words, it is possible that a few words changed overall strength of the correlation. As shown in Figure 4.6, the trends of the three languages were not so different with each other. Therefore, this difference seemed to be due to the influence of relatively low auditory or high gustatory strength for L2 English and L1 Korean rather than a cross-linguistic difference in the relationship between the senses. Other weak positive or negative correlations (Interceptive and Auditory, Haptic and Interoceptive, Olfactory and Interoceptive) were found in only one or two language groups. Since they were not robust, it was unlikely that they reflected a significant cross-linguistic differences.

4.3.6. Relationship between modalities and concreteness

The correlation between perceptual strengths and concreteness were examined to see whether concreteness reflected or summarized the perceptual properties of the words. For the concreteness of each word, the concreteness rating in Brysbaert et al. (2014) was employed. Since the survey on concreteness was not conducted for each language group, especially in Korean, this analysis should only be regarded as exploring trends. Concreteness rating had a weak negative correlation with audio and a moderate negative correlation with interoception, while having a weak correlation with olfactory, a moderate positive correlation with haptic and a strong correlation with vision (Table 4.5). Negative correlation with audio and positive correlation with vision, touch and smell were also reported in previous studies (Connell & Lynott, 2012; Vergallito et al., 2020). The result seemed to be congruent with our intuition in that something visible or touchable is generally considered more concrete. Those with high visual and haptic strength are obviously more visible and touchable, thus more concrete, while those with high auditory and interoceptive are not, so more abstract. Olfactory words are not always visible and touchable but might have some concreteness in them since experience through smell often allows imagining a physical entity related to it.

In addition to the correlation between concreteness and six modalities, the correlation with max strength and exclusivity was also analyzed. Max strength is the highest score among perceptual strengths of six modalities, which was one of the best predictors for the response times and accuracies of lexical decision tasks (Connell & Lynott, 2012; Lynott et al., 2020). Max strength was strongly correlated with concreteness. This result was expected because more than 80% of the words were visually dominant, the exclusivity of the visual words was high, and visual strengths had a strong correlation with concreteness. In fact, visual and max strength also showed a strong correlation with each other. However, the correlation with max strengths was not larger than that with visual. Exclusivity and concreteness showed a weak correlation. In other words, the less dimensional, the more concrete. Given that words with high exclusivity were mostly visual words and those with low exclusivity were gustatory or interoceptive, this also seemed to be predictable from the correlations between each modality and concreteness.

		Aud.	**Gus.**	**Hap.**	**Int.**	**Olf.**	**Vis.**	**Max.**	**Excl.**
Con.	**L1K**	-0.21***	0.01	0.33***	-0.45***	0.11***	0.48***	0.5***	0.24***
	L1E	-0.23***	0.09**	0.51***	-0.41***	0.21***	0.54***	0.57***	0.13***
	L2E	-0.15***	0.05	0.41***	-0.49***	0.13***	0.57***	0.55***	0.23***

Table 4.5 The correlation matrix between concreteness and each modality, max strength, and exclusivity. Orange shade indicates negative correlation and blue positive, with darker color being stronger correlation. *: $p<0.05$, **: $p<0.01$, ***: $p<0.001$, Aud.=Auditory, Gus.=Gustatory, Hap.=Haptic, Int.=Interoceptive, Olf.=Olfactory, Vis.=Visual, Max.=Max Strength, Excl.=Exclusivity, Con.=Concreteness

4.3.7. Individual cases

THERMAL / HARD

When there was a difference in meaning and sensory experiences related to it between languages, the difference seemed to be sensitively reflected in perceptual strengths. *Thermal* was regarded as a word related to haptic and interoceptive modality in both L1 English and L2 English, while the haptic strength was evaluated relatively low in L1 Korean (Figure 4.7-Top). This difference seemed to be because the Korean translation of *thermal*, “열의” has also a different meaning, *enthusiasm*, which is less likely to be experienced through haptic sense. It was unclear which of the two meanings is used more frequently in Korean, but given the low haptic strength, Korean participants seemed to generally understand this word as the meaning of enthusiasm. The example of *hard* mentioned above (Figure 4.7 -Bottom) was a similar example with this, but in this case English was a polysemy and one of these meanings was selected as the Korean translation (“hard” in English and “어려운 difficult” in Korean). In Korean, as the meaning of “solid” was lost, the strength of haptic was lower than in other groups.

The findings of *thermal* also demonstrated that the evaluation of individual words might differ even within a group of similar characteristics. It was interoceptive-dominant in L1 English (perceptual strength: 2.81), but it was haptic-dominant in Lynott et al. (2020). The perceptual strength of haptic sense in L1 English, 2.75, was above average, but it was much lower compared to 3.73 in Lynott et al. (2020), which also recruited native speakers of English from the same crowdsourcing web platform, Amazon Mechanical Turk. This showed that each summary statistics such as dominant modality and perceptual strengths for individual words may vary even for similar experimental groups, although the overall profile of evaluation on *thermal* was not completely different from each other (three most dominant modalities were haptic, interoceptive, and visual for both studies). Speed & Brybaert (2021) also showed that the dominant modalities of the same word changed from its previous study (Speed & Majid, 2017). Thus, the results of individual words always need to be cautiously interpreted.

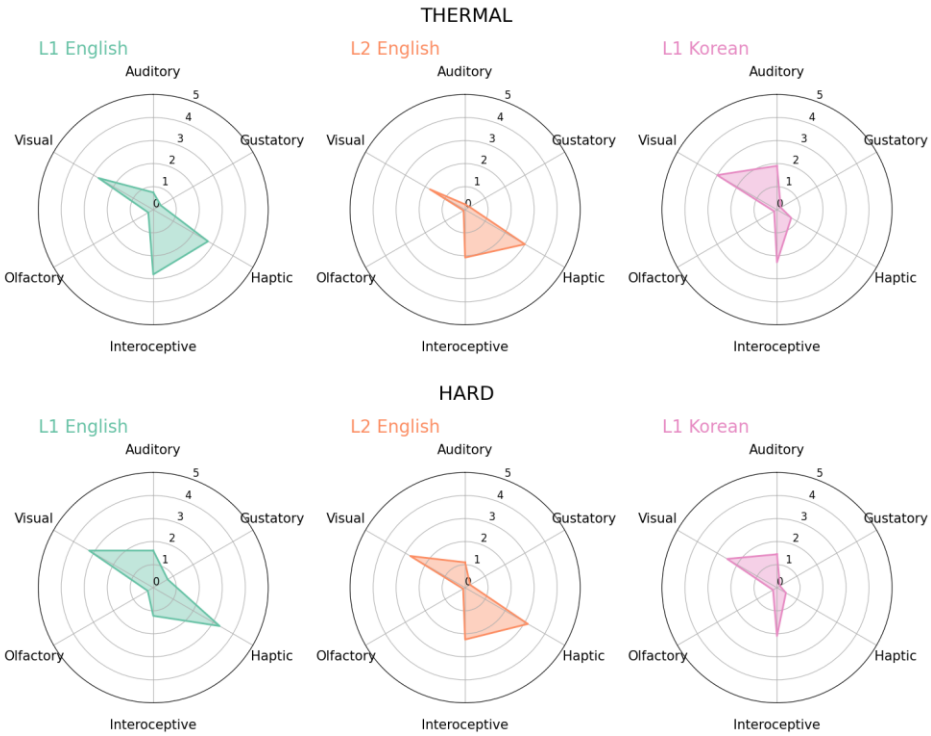


Figure 4.7. Radar charts for *thermal* and *hard*.

CALM

Some examples showed that, even when the difference in meaning between languages was more nuanced compared to that of polysemy, such subtleties can be captured by language users with the evaluation of perceptual strength. One such example was *calm* (Figure 4.8). *Calm* was interoception-dominant word in L1 English (interoception: 4.06, audio: 2.11, vision: 2.89), while audio-dominant in L2 English (interoception: 2.63, audio: 2.89, vision: 2.53) and vision-dominant in L1 Korean (interoception: 2.48, audio: 2.57, vision: 3.48). L1 English participants seemed to focus more on the (undisturbed, peaceful) internal state, while L2 English participants more on the (quiet) external environment, suggesting that L2 participants may have a different understanding or experience of *calm*. On the other hand, when *calm* is translated into a Korean word "차분한," other scores being similar, the dominance of vision increased. Although the basic dictionary definition of *clam* is related with one's inner state, it is frequently used to describe one's appearance such as in *calm hairstyle* ("차분한 머리"). In other words, "차분한" a Korean translation of *calm* had a different implication and this subtle difference was well captured in perceptual strengths and dominant modality.

Figure 4.8. Radar charts for *calm*

FEEL

An example of *feel* (Figure 4.9) was another case of subtle semantic differences between languages. Also, it showed that L2 English participants evaluated perceptual strengths under the influence of their native language, Korean. For *feel*, haptic was the dominant modality in L1 English and L2 English, but in L2 English, the perceptual strength for other senses, especially, interoceptive was as high as tactile (haptic: 3.13, interoceptive: 2.94). This may be derived from the influence of Korean. A Korean translation equivalent of *feel* "느끼다" is defined as recognition through all sensory organs rather than focusing on haptic, and is highly likely to cooccur with the experience of emotion, which is known to be relevant with interoceptive. On the other hand, in English, *feel* has a similar meaning as well, but 'touching by fingers' plays an important part in definition of *feel*, which may lead to larger exclusivity on haptic. In fact, *feel* was an interoception dominant word and its strength was 3.57, in L1 Korean. The distribution of L2 English on *feel* appeared to be somewhere in the middle between L1 Korean and L1 English. Individual cases such as *feel* suggested that, despite the overall similar distribution between languages, for certain words, the sensory experience of words may be different depending on its L1 background or usage of words.

Figure 4.9. The radar charts for *feel*. Compared to L1 English, L2 English and L1 Korean have a similar shape and occupy a wider area.

PHILOSOPHY / BIBLE

Philosophy and *bible* were cases that, as with *feel*, subtle differences in word meaning were reflected in perceptual strengths and the distribution of L2 English seemed to be influenced by Korean. However, these were different from the previous examples in that they were more related to the language users' experience on the environment surrounding words. As for *philosophy* (Figure 4.10), most of the senses were rated substantially low both in L2 English and L1 Korean. In contrast, in L1 English, it was evaluated as strong auditory words (3.88) and visual strength (3.12) was also high. Philosophy and its Korean translation “철학” have nearly the same dictionary meaning of an academic discipline or a particular set of belief. Thus, it is not the difference in dictionary meaning that had influences on differences between English and Korean. Instead, how words are actually used in everyday context was likely to create these differences in sensory experience. English speaking participants seemed to focus more on the latter meaning and associate it more to “expressing thoughts or beliefs in words or listening to them”, in that they viewed it as something that could be experienced by hearing or sight. On the other hand, Korean speaking participants seemed to emphasize the meaning of the former and regard it as something difficult to experience through the senses and unfamiliar to them. The distribution of L2 English participants was more similar to that of L1 Korean participants at this time, which might indicate that their usage in particular words was influenced more by their native language than L1 English.

Figure 4.10. The radar chart for philosophy. Compared to L1 English, L2 English and L1 Korean have a similar shape and occupy a narrower area.

Bible (Figure 4.11-top) has a conceptual or symbolic meaning, but it is also a tangible object that actually has a referent in the real world. If its meaning as an object is highlighted more, its visual or haptic properties will be more emphasized. Visual strength was scored as 4 in all three language groups. This may imply that they all regarded *bible* as an object, although it may reflect the visual imagination of biblical figures. However, haptic strength was rated as 3 or higher only in L1 English and relatively low in the other two groups. L1 English speakers have probably had more direct and indirect experiences of actually touching the Bible, so they view it as something touchable. For instance, they may have been more strongly influenced by the Christian culture and have more experience of reading the pages of the Bible or touching it for testimony. On the other hand, L1 Korean and L2 English participants perceived it as a visible, but they seemed to associate it less with haptic senses because they had little experience actually touching it.

As for *book* (Figure 4.11-bottom), a hypernym of *bible* as an object, participants of all language groups rated haptic strength strongly, although L1 English still had the highest among the languages. Therefore, it was not that L2 English and L1 Korean participants considered a book-like item itself as unbale to be experienced with haptic sense. It appeared that they regarded the Bible as rather distinct from books or had little experience touching or reading it with hand. Thus, the case of the Bible demonstrated that the perceptual strengths delicately represented the sensory experience of words and that the evaluation of L2 English participants was similar to that of L1 Korean participants with the same sensory experience.

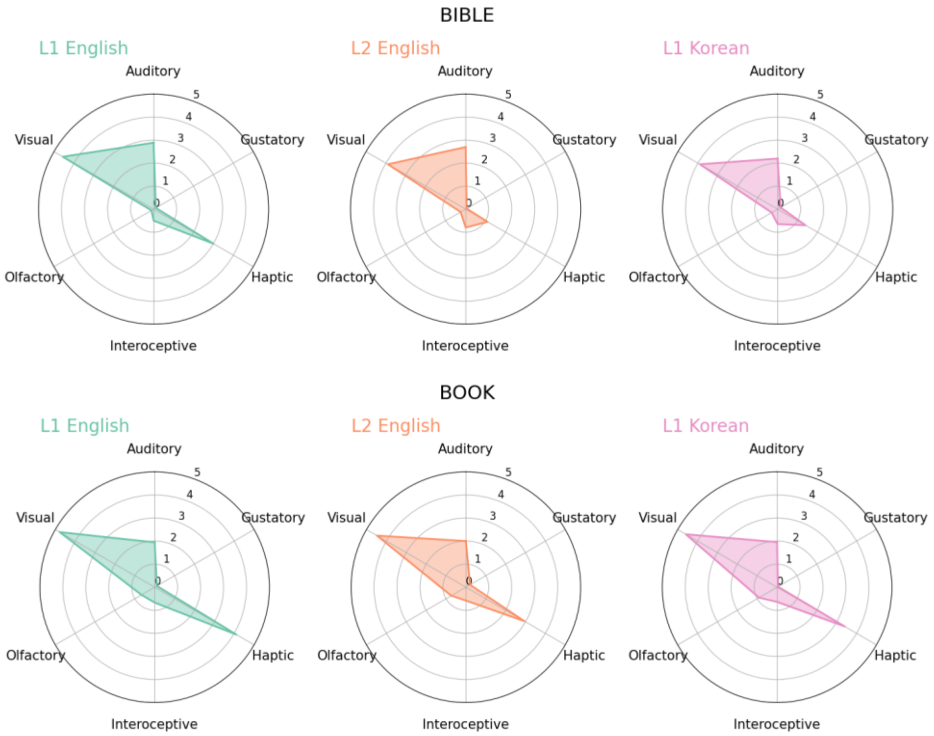


Figure 4.11. The radar charts for *bible* and *book*. For *bible*, the haptic strength of L2 English and L1 Korean is about 1, but it is 3 or higher in *book*.

4.3.8. Individual cases for loanwords

Loanwords are words borrowed from a different language (source language) and adopted by the speakers of the target language. Because the word and its referent are borrowed together, loanwords correspond with the original counterparts on the surface. However, since only a part of meanings is borrowed in the process of borrowing or the context of usage might vary, the two words often do not match completely. During the translation phase in this study, using loanwords were allowed and about 10 percent of the words in the material consisted of English loanwords. The translators were required to come up with an alternative native word before they translated English words into loanwords and select one that was more likely to be used in everyday lives; thus, the loanwords in the vocabulary list were the ones that cannot be substituted by native words or are used more frequently. Analyzing these words might provide interesting examples in the comparison of perceptual strengths between languages. On the surface, English and Korean translations correspond, so in most cases the same perceptual strengths will be observed in Korean and English. However, depending on the context or usage, they may differ between Korean and English, and for those cases, it is noteworthy examining how L2 English participants' evaluations are influenced. Despite English scripts, the words share pronunciation with Korean, so L2 English participants may perceive them to be Korean words and evaluate them similarly in Korean. Otherwise, the influence of English transcription is so strong that participants may regard the words as English rather than Korean and rate them more similarly to L1 English. This section focuses on the loanwords and offers several individual examples.

Figure 4.12. The radar charts for *cake*.

CAKE

For the vast majority of loanwords, the strengths and distribution of the three groups were similar. Typical examples were food-related terms, such as *cake*

(Figure 4.12). Since they refer to the same object and the sensory experience for them would be similar, the evaluations of the three language groups were nearly identical.

AUDIO / BLUES/ MUSICAL

L2 English participants rated their perceptual strengths similarly to L1 Korean rather than L1 English in several words. Examples were *audio*, *blues* and *musical*. Of these, *audio* (Figure 4.13) was a case where Korean loanwords had additional meaning than English. For *audio* and its loanword "오디오", visual strengths were evaluated as 2 or higher both in L2 English and L1 Korean. Since the Korean loanword also refers to an audio device, there was a possibility that "audio" can also be recognized as vision from the experience of the device. In contrast, in English, audio seldom refers to an audio device, at least by dictionary definition, so it was difficult for L1 English participants to experience it through vision. While the evaluation differed between the two groups, L2 English participants assessed it more similarly to Korean, revealing the influence from Korean.

Figure 4.13. The radar charts for *audio.*

Blues and *musical* (Figure 4.14) were cases in which only part of the

English word's meanings is borrowed. Blues refers to a genre of music or feelings of melancholy in English, while the Korean loanword “블루스” is mostly used in the former meaning and has been extended to include the meaning of dancing to such music. In English, along with auditory strength, interoception was evaluated relatively high (3 or higher in Lynott et al. (2020)), which is known to be related with emotion. On the other hand, in Korean, the word was evaluated as having relatively higher auditory and visual strength but lower interoception, reflecting that in Korean, *blues* was mainly associated only with music or dance as in the dictionary definition. The distribution of L2 English was not completely different from that of L1 English, but largely overlapped with that of L1 Korean.

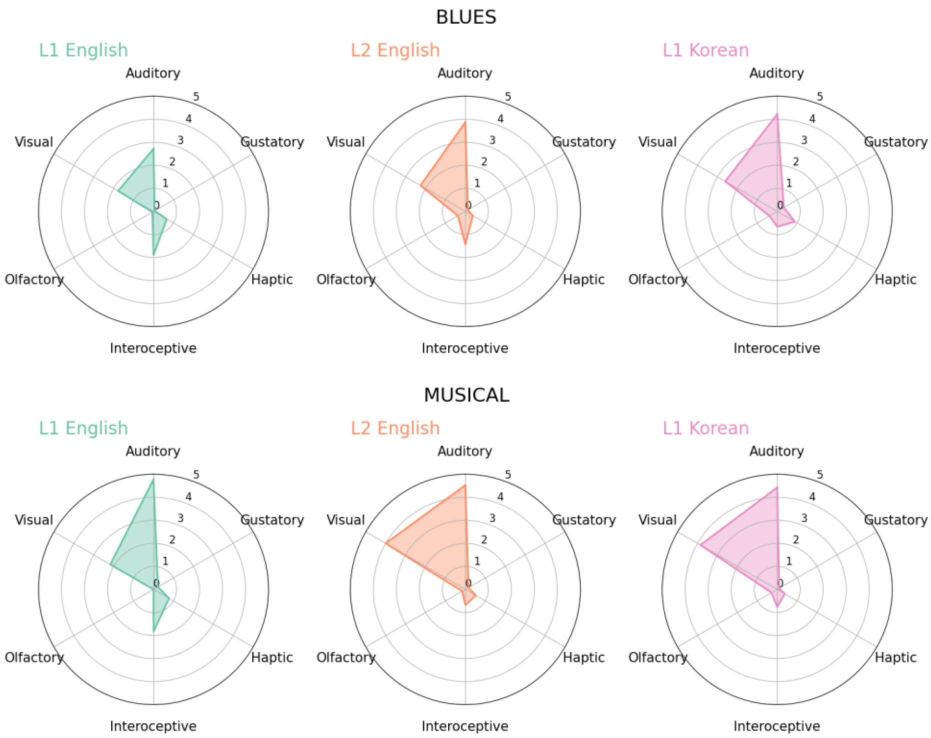


Figure 4.14. The radar charts for *blues* and *musical*.

Musical refers to an adjective of music and a type of the performance genres in English, whereas it is borrowed only as the latter meaning in Korean.

Therefore, both audio and visual strength were high in Korean, seemingly reflecting the experience of musical performance. On the other hand, the evaluation of L1 English seemed to take into account both meanings, but prioritizing the former, as shown in high auditory strength and relatively lower visual strength. The distribution of L2 English was almost identical to that of Korean, not L1 English. All of these cases demonstrated that L2 English participants recognized Korean loanwords as they would in their native language rather than English, despite the different orthographical from.

DISC

Disc (Figure 4.15) refers to rounded and flattened structures, including both optical and spinal discs. In Korean loanword, “디스크,” slipped disc was added to this meaning. Therefore, Korean speakers, encountering this word, recalled pain associated with slipped disc as well as the original meanings, resulting in relatively high interoceptive strength. On the other hand, in English, its meaning does not encompass slipped disc, so interoceptive strength was evaluated as nearly zero. Different from the examples above, L2 English participants rated this word with a distribution more similar to L1 English. However, it was unclear if this was due to the participants recognizing "disc" as an English word, rather than its Korean translation, so perceiving it similarly to L1 English participants. In Korean, the loanword "디스크" is used to refer specifically to optical discs and spinal discs or related slipped discs, rather than to rounded and flat structures in general. In fact, “디스크” in the Korean dictionary included definitions of optical discs and spinal discs, and related slipped discs respectively, but did not a definition of rounded and thin objects (National Institute of the Korean Language, n.d.). In other words, Korean speakers may treat optical discs and slipped discs as separate words, although each meaning does not seem to constitute a relationship of homonymy. In this regard, it is possibly not that the L2 participants did not connect it to the Korean translation equivalent, but that they connected it to only one of several meanings of *disc* and experienced only the senses as its connected concept.

Figure 4.15. The radar chart for *disc*.

4.4. Discussion of Experiment 1

Experiment 1 established the perceptual strengths norms of L1 Korean, L1 English, and L2 English to select the items for Experiment 2 and using the norms, explored how human sensory experiences affected word comprehension and how these were different across languages and between L1 and L2. The norms of each language showed reliable results compatible with human experience and consistent within raters and demonstrated that three language groups generally had a similar distribution in perceptual strengths despite a few differences in individual words. Further discussions of the results will be provided below, focusing on the comparison between languages, the relationship between modalities, and the relationship between concreteness and perceptual strengths.

Overall, no significant differences were found in the distribution of perceptual strengths between L1 Korean, L1 English, and L2 English. The correlation coefficients between languages were significantly large and the Euclidean distances between languages were relatively small. In other words, the similarity between languages was large. For all three language groups, visual strength was the highest, followed by auditory, haptic, interoceptive, olfactory, and gustatory. Dominant modalities also had the highest proportion of vision, accounting for almost 80% or more in all languages. However, the ranking of the most dominant modality, where vision was followed by auditory, interoceptive, gustatory, haptic and olfactory, was different from that of perceptual strength mean, which was also consistent across languages. Words with greater exclusivity were all visual words in common, and less exclusive words were generally those with high gustatory strength. It was not surprising that three groups had similar distributions, since each word used in the study referred to the same referent at least at the surface level and there was no reason to assume the environments surrounding the languages are so different from each other. In fact, the distributions in this study were also similarly observed in surveys of other languages (Lynott et al., 2020; Miceli et al., 2021; Speed & Brybaert, 2021; Vergallito et al., 2020), even though they slightly differed depending on which words were investigated—for instance, haptic strength was higher than auditory in Miceli et al. (2021), where most words were concrete nouns. However, since there was no research conducted on Korean and L2 English and no direct cross-linguistic comparison on the same word list, this comparison between L1 Korean, L1 English and L2 English provided more direct evidence for the similarity of the distribution in perceptual strengths across languages. In particular, the finding that L1 English and L2 English showed a generally similar distribution is of significance in researching L2 embodiment processing. Several studies have demonstrated that L2 processing is also embodied but their research area has been largely limited to action verbs. In contrast, this study showed that various sensory perceptions other than motor action can be experienced in a similar way and degree to that of L1. Although this was not direct evidence of L2 embodiment, it established that at least L2 words were also associated with the perceptual experience in the world.

It should be noted that participants in L1 English and in L2 English had demographic differences such as age and academic background. Participants in L2 English were a fairly homogeneous group both in terms of their academic backgrounds and ages, whereas participants in L1 English were heterogeneous in age and academic background. The environment in which a language is learned and used can have a significant impact on sensory experiences associated with language, and these demographic differences cannot be disregarded as a potential source of influence. However, a separate, unreported analysis was conducted to examine the perceptual norms of younger and older groups within the L1 English group, which revealed minor differences between the two groups, but with overall distributions that were almost identical. Moreover, it is not plausible to assume that the effect of age and academic background is greater than the difference between L1 and L2. The control of demographic characteristics is crucial for a more accurate interpretation of the influence of environment on sensory experiences associated with language. However, it does not appear that they have a decisive effect on the interpretation of the results in the current study.

The three language groups showed a similar distribution in general, but there were some differences between them. The perceptual strengths of several modalities differed between languages. All three languages were significantly different in olfactory strength, L1 Korean and L2 English in interoception, and L1 English and L2 English in gustatory. For dominant modalities, L1 Korean had a greater proportion of visual and smaller of interoceptive and haptic than the other two language groups. For exclusivity, less exclusive words in L2 English and L1 Korean were more likely to be interoceptive words, but gustatory words in L1 English. However, these findings must be interpreted with care. First of all, while certain aspects were statistically different between languages, the overall pattern was not completely distinguished from each other. For example, although the interoceptive or haptic dominant words in the other languages were labeled as visual word in L1 Korean, their interoceptive and haptic strength were higher than average. If one or more participants had rated visual strength lower, those words would not have been classified as visual dominant. Previous research found that the dominant modality of some words can be changed even when similar methods were applied to the similar groups of participants (Speed & Brybaert, 2021). Moreover, since Korean and English do not have an exact one-to-one correspondence, it is difficult to interpret the difference between Korean and English simply as caused by linguistic difference in sensory experiences. Not only polysemy or homonym but also any word translated into another language might produce subtle differences in lexical meaning from the original. For this reason, the word list of Korean and English may not be identical despite careful design of materials. Therefore, it would be more valuable to compare overall trends and distributions rather than highlight a few differences in perceptual strengths or exclusivity between languages.

Comparisons between L1 and L2 English, in contrast to the one between Korean and English, contained the same words in the word list at least on the surface, thereby reducing the possibility of inconsistencies in the lexical meaning of words between languages. The significant difference between L1 and L2 English was the average strengths of gustatory and olfactory sense. L2 English participants evaluated the strength of both senses significantly higher than L1 English participants. Since most words of gustatory and olfactory strength were extremely low in average, rated as close to zero, the evaluation of a few words having higher gustatory and olfactory strength may contribute to this overall difference. However, this did not seem to be the case, given that the number of words with high olfactory and gustatory strength was about the same between L1 and L2 English and their average strengths were rather greater in L1 English as opposed to the average strength of overall words. For instance, in both L1 and L2 English, there were 17 words with high gustatory strength (3 or higher) and the words were identical in two languages. The average strength of L1 was 4.54, which was considerably higher than the average strength of L2, 4.39. Instead, this average difference seems to arise from the difference in the number of words evaluated as zero. In L1 English, 138 words for gustatory and 128 words for olfactory were evaluated as zero for all raters, while only 11 and 8 words in L2, respectively.

Interestingly, this propensity of L2 English participants to avoid zeros was not restricted to gustatory and olfactory. Individual rating data for all senses indicate that they employed a scale between 1 and 4 more frequently, relative to other language groups (Figure 4.16). Since L1 Korean participants also appeared to exhibit this tendency compared to L1 English participants, this may be attributable, in part, to differences in cultural and academic backgrounds between Korean participants and native English speakers. However, as L2 English participants were more prone to this pattern than L1 Korean participants, consideration must be given to the characteristics of L2 processing. First, compared to speakers of native languages, L2 speakers lacked word familiarity or knowledge (Word familiarity in L2 English: 2.82; in L1 English: 2.93; L1 Korean: 2.90). When attempting to answer a question on a topic with which they were unfamiliar, ones were likely to avoid extremes since it was difficult to provide an answer with certainty. However, familiarity alone cannot explain this tendency, as the tendency persisted even within the items for which the familiarity was evaluated as 3, or “very familiar / know very well”. Another possible explanation for this is that it was caused by a qualitative or quantitative lack of L2's sensory experience of the word. To assess perceptual strength of a word, it was required to recreate the sensory experience associated with it. So, the richer the sensory experience one has of a word, the easier it will be to assess their perceptual strengths for that word. However, this experience will be rather limited in the L2 word learning environment. Even if L2 participants were familiar with the meaning of a word, in most cases, they learned it explicitly and had minimal or indirect exposure to it. Thus, L2 participants may have difficulty during the process of reproducing the sensory experience due to the lack of experiences or may have to perform indirectly through their L1 Korean assumed to be connected to L2 in their mental lexicon. This unstable process of representation may result in choosing a scale in a way that leaves room rather than entirely eliminating the possibility of the word being experienced with any modality.

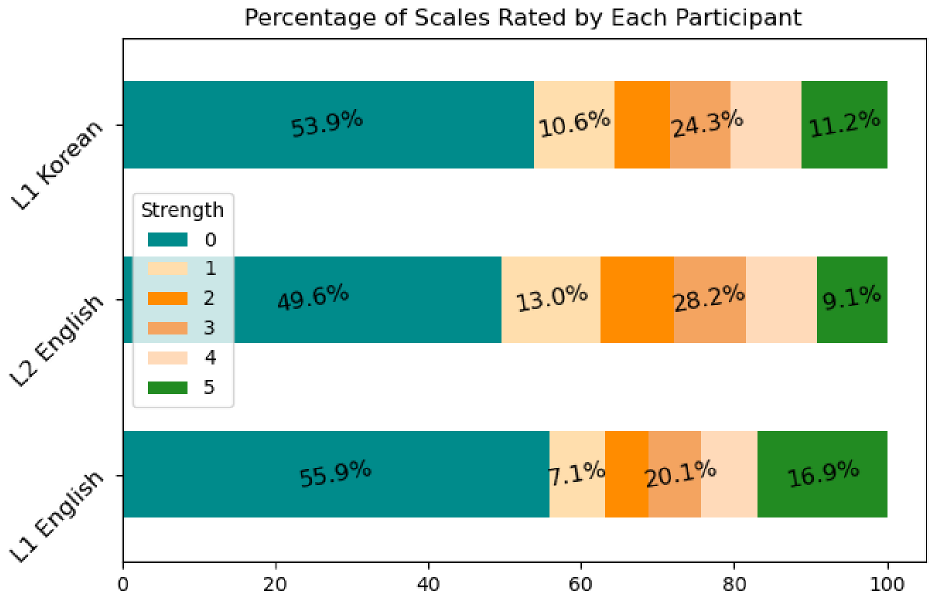


Figure 4.16. Stacked percentage bar graphs displaying how frequently each scale was used by participants. L2 English participants tended to avoid scales at both ends, such as 0 and 5, compared to L1 English group. Considering that the two groups were heterogeneous with different cultural and academic backgrounds, this can be seen as a difference in the response style between the two groups. However, even though L1 Korean and L2 English participants were relatively homogenous groups, they showed differences in how they dealt with extreme values. In particular, L2 English participants have an inclination to avoid using zeros, seemingly using 1 instead of 0. The sum of frequencies of 0 and 1 were similar in all three groups as 52-54 percent.

Although comparisons across languages based on differences in average strengths or exclusivity were often difficult to interpret due to diverse factors mentioned above, differences in individual words might be able to provide some insights on how differences between languages are reflected in perceptual strengths. The cases of *feel* and *calm* explained in the result section were a few of these examples. The subtle variation in meaning between Korean and English words seemed to be mirrored in their perceptual strengths. For instance, the Korean translation of *feel* relates to emotions and a broader range of senses than its English counterpart, which was revealed in the high interoceptive strength and low exclusivity in the Korean rating. The perceptual strengths of L1 Korean and L1 English captured the variation in meaning between English word and its Korean translation with great sensitivity. When there was no one-to-one relationship between the English and Korean translations for a particular word, the rating of the word became considerably different. For example, if an English polysemy loses some of its meanings when translated into Korean, only the perception of one meaning was evaluated in Korean (e.g. *hard* was not rated as haptic in Korean because it is translated into "difficult"). This sensitivity demonstrated that participants did not judge the perceptual strength of language randomly, but rather in accordance with genuine language perception. Obviously, this sensitivity, which led to an intriguing analysis, is also a characteristic that makes cross-language research more challenging. Careful selection of items is essential when comparing the same word in different languages as in Experiment 1 and 2.

An interesting question to consider while analyzing individual words was whether the L2 English participants' ratings showed a distribution more similar to L1 English or L1 Korean when the distributions between them were distinct. An analysis of this will show how linguistic or cultural factors affect the assessment of perceptual strengths. However, because Korean and English do not completely correspond, not all words are proper for this. One possible type of words for this analysis is those in agreement in translation. For example, *eagle* was translated identically as "독수리" in Korean by all translators. Even if the usage and experience of these words vary between languages, they are at least matching in form. When processing these words, most of L2 participants would perhaps represent the similar distribution with the Korean translation. *Feel* was an example of such words. As for *feel*, L2 English participants appeared to make more comparable judgements to L1 Korean than L1 English. As shown in the radar chart (Figure 4.9) its form of the graph was more akin to that of L1 Korean and its Euclidean distance was closer to L1 Korean (L1 English vs L2 English: 2.51, L1 Korean vs L2 English: 0.96). Another possible type of words was loanwords, because the source and target words, at least on the surface level, referred to the same referent. Comparisons in loanwords showed that words referring to the same object were in general rated with similar perceptual strengths. However, there were differences between Korean and English in some cases, and with the exception of few cases, the distribution of L2 English evaluation was more comparable to that of L1 Korean as in *bible* or *musical*. Not only in the individual cases but also in the overall trend, L2 English appeared to be closer to L1 Korean than L1 English. For example, the mean of the Euclidean distance between L2 English and L1 Korean was shorter (L1 English vs L2 English: 1.22, L1 Korean vs L2 English: 0.97, L1 English vs L1 Korean: 1.38). These findings might imply that L2 participants' ability to recreate their experiences with L2 word was often strongly influenced by or mediated through their native language.

The relationship between modalities did not differ significantly between languages. Although there seemed to be some differences (for example, negative correlation between auditory and gustatory or olfactory was found only in L1 English), since the majority of the differences were weak correlations and all languages exhibited a generally similar pattern, it was difficult to conclude that the differences were due to language differences. The correlation between modalities was congruent with the intuition on human perception experience. A positive correlation between olfactory and gustatory, positive correlation between visual and haptic, negative correlation between auditory and visual, and negative correlation between interoceptive and visual were all consistent with our understanding of human perception and they were also observed in previous studies. As indicated by the excellent inter-raters' consistency, the raters' assessments were not arbitrary but consistent among themselves and this well-agreed evaluation seemed to represent the experience of human perception. Admittedly, this does not prove that embodiment is essential for the comprehension of words, but it does demonstrate that at least it is possible for language users to plausibly reproduce the experience of the world with language as a cue; that is, language and the perceptual experience of the associated referent interact with each other.

One of the observed correlations worth noting was the negative correlation between interoception and vision. Because interoception has been rarely explored and included in perceptual strengths as one of modalities; hence, investigations on their correlation with other modalities are also uncommon (Lynott et al., 2020). This study confirmed, consistent with earlier research, that interoception was negatively correlated with vision in Korean and L1 and L2 English. The finding that interoception is negatively correlated with vision suggests that interoception should be addressed in perceptual strength research alongside the five senses. Interoception can compensate for the aspects of the world that cannot be explained by vision, the sense that accounts for the biggest portion of the sensory modalities. When perceptual strength is assessed without interoception, a word with weak visual strength may be deemed incomprehensible by sensory experience.

The importance of interoception was also revealed in the relationship between concreteness and modalities. Concreteness had a positive correlation with visual, haptic, and olfactory, and a negative correlation with interoceptive and auditory. Interoception was only modality having a modest negative correlation with concreteness. If the sensory experience of abstract words is evaluated without interoception, the majority will be rated as having low perceptual strength in any sense and will be considered as words that cannot be experienced via the sensory modalities. Although abstract words are sometimes evaluated as being related to auditory modality, audio alone will not be able to describe all of the sensory experiences associated with abstract words. Evaluating abstracts words not having any sensory experience may lead to the erroneous conclusion that the embodiment of abstract concepts cannot be explained only with the sensory-motor domain. Therefore, constructing perceptual strength norms including interoception is necessary to explain abstract words as well through embodiment cognition theory.

It is also worth noting that the correlations between concreteness and each modality varied. Some were negatively correlated, some were only weakly correlated, and some were not, such as gustatory. In other words, it was difficult to capture all the senses with a single measure of concreteness. Perceptual strength is likely to measure some different properties than concreteness or imageability, perhaps something in a broader range. Therefore, it is difficult to assess the words encountered with such various senses only through the abstract-concrete dichotomy.

Nevertheless, one may argue that it is more effective to employ a single dichotomy, such as concreteness, rather than multiple sensory indices. Moreover, concreteness has a high correlation with max strength, which is supposed to function as a composite of the strengths of all senses. In fact, previous studies have shown that concreteness is highly related with a variety of linguistic phenomena such as concreteness effect and the response times of lexical decision tasks. However, being able to predict the several linguistic phenomena only with concreteness is likely to be an illusion caused by vision dominance. Since there is a strong positive correlation between vision and concreteness, employing concreteness has the similar effect as utilizing visual perceptual strength. The visual strength alone, since most words are visually dominant, can explain large proportion of language, and concreteness take advantage of this. In this respect, the explanatory power of concreteness may be proportional to how many visual words are contained in the item list. Indeed, the well-known concreteness effect, in which concrete words are processed faster than abstract words, was reversed when emotion-related words, presumably less visual dominant words, were used as materials (Kousta et al., 2011). Therefore, substituting multiple sensory indices with concreteness may appear to be an easy solution, but it is not the most effective way to describe actual linguistic phenomena.

Chapter 5. Experiment 2

5.1. Introduction

In Experiment 2, based on the multilingual perceptual strength norms established in Experiment 1, whether this classification reflects neural activities will be explored by means of EEG experiments.

The results of Experiment 1, first, showed that language users were able to classify words based on each sensory modality, the distinction among six senses for each word seemed to be congruent with human perception, and the assessments of the raters were reliably consistent. Second, the distributions of perceptual strengths were not distinct between L1 English, L2 English, and L1 Korean, although the analysis on individual words showed minor differences between language groups, along with the cases indicating that such differences were influenced by the context of language use. All three groups shared similar distribution of perceptual strengths where vision dominates among six sensory modalities, while the proportions of olfactory and gustatory sense were small, and the dominant modality of individual words did not vary in most cases among languages. Lastly, in all language groups, interoceptive strength and concreteness revealed negative correlation, in part supporting the assumption that abstract concepts are embodied through modality of interoception.

These findings, however, were based on offline-tasks, underpinning the embodiment cognition theory only to a limited extent. As pointed out by Villani et al. (2019), this classification has to be further validated by behavioral and neurophysiological experiments. Moreover, L2 learners often reveal significant differences between their knowledge and performance during L2 processing (Jiang, 2004; 2007). For example, even though they have grammatical knowledge on their L2 English such as the third person singular -*s*, displaying almost no grammatical deficiency on test results during offline tasks, their online processing on number agreement is relatively weak or shallow, even resulting in different EEG responses from that of native speakers. Therefore, the findings of Experiment 1, that words can be categorized according to each sensory modality and differences between L1 and L2 is negligible, requires further evidence from online processing, particularly the one demonstrating the neural activities of human brain. Experiment 2 will investigate whether linguistic concepts of each sense are processed differently in the modality-specific regions of brain and how their embodiment differs between L1 and L2.

5.2. Previous studies

Many previous studies have attempted to support the claims of embodiment cognition theory by exploring feature-specific neural responses (Bergen et al., 2010; Cervetto et al., 2021; Dalla Volta et al., 2014; Hauk, et al. 2004; Hauk & Pulvermüller, 2004; Pulvermüller et al., 2009; Tettamanti et al., 2005). From the perspective of embodiment cognition theory, vocabularies with different sensory characteristics will activate different regions of brain, which will be reflected in different behavioral and neural responses. Numerous behavioral and neuropsychological experiments have examined language processors' responses across the different types of features including body effectors, modalities, and abstract classification.

Each body effector, Bergen et al. (2010) showed, activates effector-specific sensory-motor circuits through a series of behavioral experiments. In a verb–verb near-synonym task using the priming technique, when presented two words (‘run’-‘kick’) of the same body effector (leg-leg) as prime and probe word, participants took longer to judge a pair of words as non-synonym, compared to when the words (‘run’-‘drink’) for different body-effectors (leg-mouth) were given. This interfering effect might suggest that two actions employing the same effector inhibit each other within the same neural circuit, while the ones with different effectors are represented in their respective domains, not interfering with each other. This effector-specific response was further supported by several experiments using EEG and fMRI (Cervetto et al., 2021; Dalla Volta et al., 2014; Hauk, et al. 2004; Hauk & Pulvermüller, 2004; Pulvermüller et al., 2009; Tettamanti et al., 2005). Dalla Volta et al., (2014) found that when participants performed a sematic decision task where they decided whether the presented word was concrete or abstract, their EEG response elicited the involvement of motor cortex for action verbs stimuli in contrast to abstract verbs, and this activation significantly differed according to implied body effectors. For example, hand verbs activated a ventro-lateral area of the right premotor region, whereas a dorso-medial part of premotor areas was bilaterally involved for foot verbs.

Feature-specific neural responses across different modalities such as action, sound and vision has been investigated in many studies (Harpaintner, Sim, et al., 2020; Harpaintner et al., 2020; Hoenig et al., 2008; Kiefer et al., 2008; Kiefer et al., 2012; Kuhnke et al., 2020; Popp et al., 2016; Trumpp et al., 2013; Trumpp et al., 2014; quotation). Trumpp et al. (2014) demonstrated unconscious feature-specific EEG responses between action and sound nouns, using masked repetition priming technique which is suitable for examining restricted lexical processing. In the experiment, participants were given a task to silently read the word pairs of action-related and sound-related concepts. Each group of word pairs was divided into non-repetition and repetition conditions, where prime and probe words were identical. It is known that repeated stimuli evoked faster and more accurate responses (Adelman et al., 2014) and characteristic ERP components such as N1-P2, N400, and often LPC (late positive component) (Kiefer, 2005; Misra & Holcomb, 2003; Schweinberger et al., 1995). The results showed that priming effects varied between sound and action words in terms of polarity, topography, timing, and neural generator. While sound words elicited more positive potentials in both left and right hemisphere with temporal and other sources from 250 msec time window, action words evoked more negative potentials only in left hemisphere with neural generators of left frontal and parietal cortex from slightly later time window. These results suggested that the ERP responses varied depending on modalities from the early time window under limited lexical processing. Similar distinct ERP modulations were obtained also for action and sound verbs in the lexical decision task (Popp et al., 2016). This indicated modality-specific activation of brain was not limited to a certain grammatical element such as nouns, but also appeared in others such as verbs.

Several studies have attempted to investigate feature-specific neural responses across modalities including abstract concepts (Catricalà et al., 2014; Dalla Volta et al., 2014; Dreyer & Pulvermüller, 2018; Harpaintner et al., 2020; Harpaintner, Sim et al., 2020; Muraki et al., 2020; Villani et al., 2021; Vigliocco et al., 2014; Wilson-Mendenhall et al., 2013; Zhang et al., 2020). Since there has been no agreement on how to categorize abstract concepts in terms of embodiment, each study classified abstract concepts in its own way and measured neural responses with that criterion. One group of studies found the characteristic neural response of abstract concepts by contrasting them with concrete concepts (Catricalà et al., 2014; Dalla Volta et al., 2014; Dreyer & Pulvermüller, 2018; Vigliocco et al., 2014). For example, in Dalla Volta et al. (2014), abstract words elicited mostly frontal scalp activation whereas concrete words evoked parieto-frontal potentials depending on body effector. These studies either claim that abstract concepts are not processed in sensory-motor areas, in contrast to concrete concepts (Dalla Volta et al., 2014; Vigliocco et al., 2014), or that they can be treated in the same way with only a difference in degree from concrete concepts (Catricalà et al., 2014; Dreyer & Pulvermüller, 2018). Another group of studies, taking into account the heterogeneity of abstract concepts, classified them with criteria suitable for the characteristics of abstract concepts (e.g. emotional, social, mental, and numerical aspect) and obtained the EEG response accordingly (Catricalà et al., 2014; Dreyer & Pulvermüller, 2018; Muraki et al., 2020; Vermeulen et al., 2007; Villani et al., 2021; Wilson-Mendenhall et al., 2013; Zhang et al., 2020). This type of research assumed that abstract-specific regions, such as those related with emotional or social valence, are involved for processing abstract concepts. Muraki et al. (2020) classified verbs into four types—abstract mental, abstract emotional, abstract nonbodily, and concrete—and compared ERP waveforms between them. The results showed that some types of abstract verbs elicited distinct polarity, topography, and neural generators, even though each type did not have a unique representation for its own. For example, in comparison to abstract mental state and concrete verbs, abstract nonbodily state verbs displayed a prolonged negativity at frontocentral region and a sustained positivity at parietal and occipital region after

400 msec time window. The other group of studies examines modality-specific neural patterns of abstract concepts, presuming that abstract concept is activated in the sensory-motor area as well (Harpaintner et al., 2020; Harpaintner, Sim et al., 2020; Villani et al., 2021). These studies consider abstract concepts to be processed by the similar embodied mechanism with concrete concepts, so rather than dichotomously dividing words as abstract or concrete, the studies classify abstract concepts similarly as concrete concepts are divided by sensory modality or body-effector and examine their neural processing. In Harpaintner et al., (2020) (see also Harpaintner, Sim, et al. (2020) for an fMRI study), ERP responses to abstract concepts associated with vision or action were compared. Both in a lexical decision task and a conceptual decision task, where participants determines whether two words were semantically related or not, feature-specific ERP modulations were found in early and late time windows.

However, although many studies have shown that different characteristics of body effectors, modalities and abstract concepts elicit different behavioral and neural responses, there are only a few studies dealing with feature-specific processing of L2 learners (Bergen et al., 2010; Vukovic, 2013; Zhao et al., 2020). Those studies have demonstrated that similar feature-specific responses were observed in L2 learners. For instance, in Bergen et al. (2010), L2 learners showed interference effect within the same body-effector as L1 speakers did and this interference effect was correlated with learners' proficiency. Zhao et al. (2020) demonstrated the interference effect of sensory modalities. L2 learners' responses were slower when vision and audio were switched, due to the interference of modality switching. The effect resulted from both comprehension (lexical decision task) and production (word-naming task), and did not disappear even by adding a set of fillers between the targets to attenuate the activation level of the perceptual modality. This interference effect was also affected by L2 proficiency, and in general, the effect tends to increase when L2 proficiency was high. As seen above, most of previous studies on L2 processing conducted behavioral experiments. It is difficult to find L2 research on the neural responses of feature-specific processing. Even if there are several neuropsychological studies on L2 embodiment processing, in most cases, they dealt with motor resonance in which comparison between action and non-action language were a main focus (De Grauwe et al., 2014; Vukovic, 2013; Vukovic & Shtyrov, 2014) or word acquisition using a word recall test (Kelly et al., 2009).

Previous studies have provided evidence for the embodiment of language through the use of behavioral, EEG, and fMRI experiments, demonstrating feature-specific responses. However, much of this research has focused on body-effectors or action verbs, which are among the most commonly studied topics in embodiment theory, and even research on sensory areas has largely been limited to visual and auditory modalities. Furthermore, there have been few studies that have applied this methodology to L2 processing, and to the best of my knowledge, no study has investigated L2 learners' neural responses across sensory modalities.

In this regard, the present study explores the feature-specific neural responses with a focus on abstract concepts in L2 processing and compares them with those from L1 processing. To this end, this study aims to explore feature-specific neural responses, including interoception, which has been rarely researched. Interoception is one of the major human sensory modalities but is often disregarded from previous studies. Interoception is known to play an important role in the understanding of abstract concepts including emotional words (Connell et al., 2018). This study includes interoception as a crucial modality for understanding abstract concepts, considering that the embodiment of abstract words can be processed in sensory-motor areas similarly as are concrete words.

5.3. Present study

Experiment 2 examined whether modality-specific neural responses between audio and interoception occur in both L1 and L2 processing, using a masked-repetition prime paradigm. In the experiment, participants were instructed to perform a lexical decision task where audio and interoception words were displayed. The task was designed as a within-subject design where each participant was presented both with their L1 Korean and L2 English. During the task, their response times, accuracy and EEG responses at the target words were recorded. With EEG data, ERP analysis, time-frequency analysis and source estimation analysis were conducted.

In this study, the embodiment of abstract concepts will be investigated by comparing responses across audio and interoception sensory modalities. Interoception, which is an important but often overlooked sense for perceptual information, is particularly relevant to the exploration of abstract concepts. Despite recognition of interoception as a key aspect of embodiment of abstract concepts in L1 research, there have been relatively few studies that focus specifically on this modality. Given the heterogeneity of abstract concepts, it is unlikely that interoception alone can fully account for embodiment, but a comprehensive examination of interoception may aid in our understanding of these concepts. Audio was selected as the modality to be compared with interoception, rather than vision, which is the most prominent modality. This is because the perceptual dominance of vision may be too strong in both L1 and L2 to allow for the detection of other influences, and visual presentation of the stimulus is inevitable in the task. On the other hand, an auditory medium is not employed in the task, and according to the results of Experiment 1, audio and interoception have similar perceptual dominance. Additionally, previous research has already demonstrated audio-specific responses, and the typical neural patterns associated with audio are relatively well-known, which will aid in ensuring the validity of the experiment.

Previous studies utilized diverse tasks such as a silent reading task, lexical decision task, semantic decision task, and syntactic decision task. A silent reading task and lexical decision task are more implicit than the other two tasks. A sematic decision task (or conceptual decision task/concept judgment task) and syntactic decision task directly ask the participants to associate with target concept or modality (*e.g.* whether the word is concrete or abstract / sound-related or action-related and whether the pair of words has similar meaning / falls in the same category.), whereas a silent reading task or a lexical decision task only requires understanding words at implicit level. In this study, the implicit tasks are preferred, because they reduce strategic access to the semantic feature of a word, giving only an implicit access to the meaning of a word, which makes them more suitable for investigating unconscious processing. Although several studies showed that activation of feature-specific regions occurs regardless of the tasks (Dalla Volta et al., 2018; Harpaintner et al., 2020; Kuhnke et al., 2020), a few studies have investigated the task effect on L2 learners, which might work differently as the off-line and online-tasks induced different responses for L2 learners. Among the implicit tasks, a lexical decision task is favored over a silent-reading task because it demands more attention from participants and provides information on their word knowledge with accuracy rate. Given that L2 learners' proficiency or their word knowledge might affect the results (Bergen et al., 2010), it is necessary to induce participants' concentration and measure their level of knowledge by excluding incorrect answers.

This study methodologically employs a masked repetition priming technique in order to explore restricted and automatic word processing. It is a refined method to unravel the underlying mechanism of the cognitive process (Forster et al., 2003). There are various versions of this technique, but a typical method is to place a *forward mask* and a *backward mask* (e.g. a series of hashtags or random letters) before and after a *prime* that lasts for a short period of time (e.g. 40 msec), followed by a *target*, which is identical with a prime (Van den Bussche et al., 2009). Although meaning of a prime, or even its presence, is not perceived by participants due to short duration and masks, this short exposure before a target has a great impact on their ensuing processing, which is often referred to as the repetition effect.

At a behavioral level, the repetition facilitates participants' processing. They would respond to a repeated target faster and more accurately compared to a non-repeated target or neutral baseline (Forster et al., 2003). At a neural level, while the repetition reduces neural activity in most instances ("repetition suppression") (Grill-Spector, Henson & Martin, 2006), under certain circumstances it increases neural activity depending on the type and number of repetition ("repetition enhancement") (Müller et al., 2013; Turk-Browne et al., 2007). For instance, when a stimulus is provided with a low degree of visibility (Turk-Browne et al., 2007) or when the number of repetitions is low (Müller et al., 2013), repetition enhances brain responses. In ERP studies, it is known that repetition evokes distinctive ERP modulations such as N1-P2, N400 and LPC (Kiefer, 2005; Misra & Holcomb, 2003; Schweinberger et al., 1995). By using the repetition effect to track the neural activities that are enhanced or reduced, it is possible to explore the neural activities required for specific processing of interest.

Although the repetition effect is equally applied to both masked priming task and unmasked priming task in which a prime is visible, in this study, a masked repetition priming task is adopted, because masked words activate more restricted areas of brain, compared to unmasked words. Although this characteristic appears to be a disadvantage at first glance, it is in fact a useful trait when studying the activation of a certain region at an early stage of processing. When information is processed in our brain, it is immediately transmitted to the most parts of brain, which often makes it difficult to investigate the area at issue (Foster et al., 2003). In the research of embodiment cognition, it is necessary to disambiguate whether embodiment is essential for the understanding of language or is a secondary process following the initial semantic access. If a certain neural response automatically and subconsciously occurs in a limited way at the early stage of cognitive process, it is likely not a secondary process. The masked repetition priming paradigm provides the possibility of observing such neural responses. If distinct modality specific responses are found in the repetition effect induced by the masked prime, it may indicate that modality-specific processing is present even in the unconscious and automatic process.

Although in L2 research, a between-subject design in which the same language is compared between two monolingual groups is often used, this study was designed as a within-participants design. Both between-subject design and within-subject design have been employed in the L2 embodiment cognition research, but within-design should be preferred in that it can reduce the influence of socio-cultural experience, which is likely to interact with language representation (Monaco et al., 2019). In fact, neural representations of different languages within a single bilingual group may overlap greater than in two monolingual groups (Xu et al., 2021). Furthermore, as demonstrated in Experiment 1, L2 speakers' interpretation of words is occasionally impacted by their native language, resulting in their perceiving the same word differently from L1 speakers. Therefore, within-subject design is likely to be appropriate to focus on linguistic factors other than socio-cultural ones.

5.4. Research questions and predictions

5.4.1. Research questions and general predictions

Based upon the assumptions of embodiment cognition theory and the findings of Experiment 1 and other previous studies, the following predictions can be made. Theoretically, it was assumed that language processing is modality-specific in each sensory-motor domain and in Experiment 1, the participants evidently distinguished the modality of each word. In this regard, distinguished neural responses will be found from each sensory group. In other words, different neural responses might be elicited depending on whether participants read audio-dominant words or interoception-dominant words. Their EEG responses will not only reveal a variation in waveforms and powers in certain frequency bands, but their source of neural activities will be estimated in specific sensory-motor area where each modality is processed. Moreover, if embodiment processing is not merely a by-product of language processing, the modality specific responses will also be reflected in the repetition effect. For example, a larger or smaller repetition effect may occur depending on the modality. Meanwhile, if this processing differs for each language, differences in EEG responses will be observed depending on whether participants read the word with their L1 or L2. Since Experiment 1 found the similar distribution of perceptual strengths between L1 and L2, it is first predicted that there would be no significant difference in EEG responses across language groups. However, because it is also assumed that the degree of embodiment is smaller in L2 (Bergen et al., 2010; Vukovic & Shtyrov, 2014), less clear or no distinction was observed between modalities from ERP waveforms or source analysis in L2 data. In particular, the neural responses of abstract words or interoceptive words related to abstractness may reveal a greater difference between L1 and L2, for abstract words are known to be acquired later with more difficulty (Ghio et al., 2013; Fini et al, 2021).

5.4.2. Related neural indicators and predictions

As an exploratory study, this study will consider a broad range of possible results rather than restricting modality-specific responses to several selected neural indices. In fact, since no previous research has compared the same types of materials or features as this study, the observed neural responses cannot be directly compared with the findings from the previous studies. However, there are neural indicators predominantly employed in neurolinguistic research and are presumed to be relevant to the results of this study, so the present study will primarily attempt to interpret the results focusing on them. (For details on ERP components, frequency bands and brain regions of interest, see *3. Neurological background* section.)

5.4.2.1. ERP components

P2 (N1-P2 Complex)**:** P2 (or P200) is associated with early perceptual processing and other cognitive activities in the time window of 100-250 msec. It is a component of interest in this study because it is related to repetition effect or semantic factors affecting early lexical processing such as emotionality. In terms of modality-specific neural responses, interoception is more related to emotionality than audio, so a larger P2 may be found in the interoception condition. However, prediction is not straightforward because P2 is influenced by a variety of complex cognitive factors.

P3 (P300)**:** The amplitude and peak latency of the P3 component may be influenced by the attentional allocation mediated by the overall level of arousal of the processing system. With respect to the arousal level, the difference in amplitude between audio and interoception can be interpreted as reflecting the difference in overall arousal level during word processing. For example, if processing audio words is characterized by a relatively lower arousal level compared to interoception, it will result in a relatively higher amplitude. Given that P3 is associated with a wide range of cognitive processing, it is difficult to predict how it will be influenced by modalities. One possible factor that can affect the difference in cognitive arousal level between the two modalities is emotionality. However, the effects of emotionality on P3 in previous studies have been mixed. While some studies have found that high emotional valence, particularly negative emotional words, elicit smaller P3 amplitudes (Meinhardt & Pekrun, 2003; Schupp et al., 2006), others have demonstrated larger P3 for high emotional words (Herbert et al., 2006; Kanske & Kotz, 2007).

N400: Since the N400 is involved in a variety of verbal or non-verbal semantic processing, numerous factors can affect modality processing. If audio and interoception have different neural generators, they will have different access to semantic memory, resulting in a variation in the N400 amplitude. In addition, concreteness and emotion are recognized as characteristics associated with N400. Abstract words and emotional words generally elicited smaller N400 compared to concrete words and neutral words, respectively (Barber et al., 2013; Kanske & Kotz, 2007; Wang et al., 2019). According to these results, it is predicted that interoception conditions that are more abstract and more related with emotion will show smaller negativity.

5.4.2.2. Frequency bands

Gamma band: Gamma band oscillation reflects the neuronal interaction of sensory-perception areas. If audio and interoception words require activation of different sensory regions, the power and topography of the gamma band can be distinct. Due to the lack of prior research on audio and interoception, it is difficult to predict changes in the gamma oscillation of the two modalities. One possible prediction is that since gamma power increases in response to audio perception (Leicht et al, 2021; Morillon et al., 2010; Pantev, 1995; Steinmann et al., 2014), audio words may elicit larger gamma oscillation than interoception. However, prediction is not so straightforward because the senses related to gamma frequency include not only auditory but also visual, somatosensory, tactile and olfactory (Karns & Knight, 2009; Mori et al., 2013; Zhang et al., 2012). However, the difference in power the gamma band will at least reveal that there is a difference in neural connection of sensory-perception.

5.4.2.3. Brain regions of interest

The regions relevant to this study are auditory cortex (BA 41, 42, 21, 22), insular cortex and its related regions (BA 1, 2, 3, 24, 32, 33, 44, 45, 47,), Broca's (BA 44, 45) and Wernicke's areas (BA 22), anterior temporal lobe (BA 20, 21, 22, 38), and medial temporal lobe (BA 28, 34, 35) (For a detailed description of each area, see *3.5.2. Brain regions of interest*).

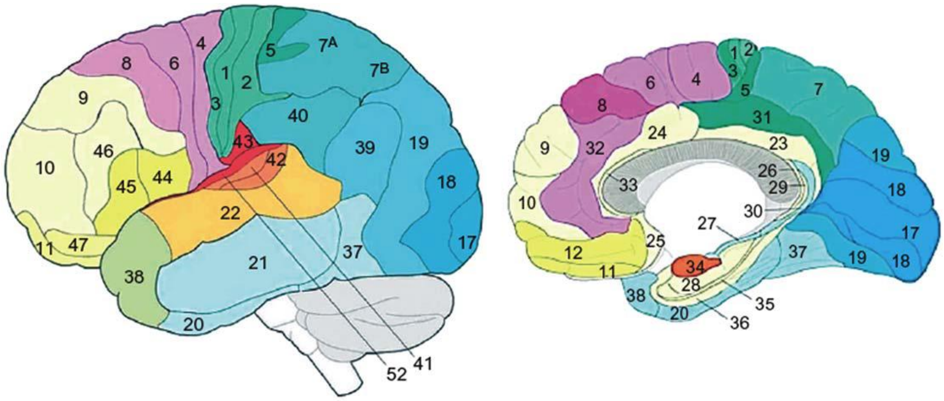


Figure 5.1. Brodmann areas. [Left] Lateral view of left hemisphere. [Right] Medial view of right hemisphere. Adapted from Gage & Baars (2018).

5.5. Method

5.5.1. Participants

35 right-handed Korean learners of English (male: 18, female: 17) were recruited for Experiment 2 from Seoul National University. The data of four participants were eliminated from the analysis for the homogeneity of the participants in terms of their L2 usage frequency, L2 learning background and

proficiency. Three of them were excluded because they reported to use English much more frequently than other participants (more than 20 hours per week), while the other participant was excluded because the self-reported English proficiency was substantially lower than the others (1.75; mean rating: 3.87). In addition, the data from one participant were also excluded due to device malfunction (the disconnection of the response pad). As a result, a total of 30 participants were analyzed. None of them participated in Experiment 1. All they had normal vision and reported no history for neurological or psychiatric disorder. Their average age was 23.47 (*sd=3.54, max=36, min=18*).

The participants were highly proficient in English as L2, according to the classification of standardized test. Their average TEPS score was 474.93 (*sd=59.01, max=597, min=387*), which is considered near-native level according to TEPS score classification. They also self-reported their proficiency score based on CEFR self-report.[8] The average level for listening was B1–B2, reading B2–C1, speaking B1-B2, and writing B2.[9] According to the CEFR classification, B1 is intermediate, B2 is upper-mediate, and C1 is expert level. By this standard, participants were upper-mediate level learners, and according to the reading proficiency considered most relevant to this experiment, they were upper-mediate to expert level learners.[10]

---

[8] Participants read CEFR's Self-assessment Grids translated into Korean and evaluated their own proficiency. (https://www.coe.int/en/web/portfolio/self-assessment-grid)

[9] A1, the lowest level, was given a score of 1, and the highest C2 was given 6. The averages of each area were 3.63 (listening), 4.31 (reading), 3.45 (speaking), and 4.03 (writing).

[10] Since CEFR was written for European learners, it may be difficult to evaluate Korean learners with the same criteria. This is because cultural influence cannot be excluded due to the nature of self-report that is affected by subjectivity. The correlation between the standardized test TEPS score and the average of the four areas of self-assessment was 0.40, which indicates moderate correlation (with listening: 0.47, reading: 0.39, speaking: 0.25, writing: 0.26).

Their first exposure to English was at an average age of 6.90 (*sd=1.76, max=12, min=4*), and most of their first exposure was through regular classes such as schools or other institutions (86.67%; through media: 10.00%, through family or friends: 3.33%). Eight of the participants had experience of living in English-speaking countries (USA: 7, Philippines: 1). All of them lived there for less than a year. Four of them lived before the age of 15 and received formal pre-elementary or elementary education. 11 participants had immersion education at a university or other institutions. Those who received immersion education in Korea had an average of 2.3 years of experience of it, and the amount of time for immersion classes was generally less than two hours per week. They learned English in a various way such as regular classes or conversation classes from schools and other institutions[11], media and books or communication with family and friends and most of them had experience of learning English in two or more ways. The most frequently used method among them was regular classes from schools or other institutions (70.00%; through books or media: 16.67%; through conversation classes from schools or other institutions: 13.33%). All participants received regular classes from schools or other institutions, with about 77 percent of them taking classes in Korean and 23 percent in English or both languages. Summarizing the participants' English education background, the majority of them learned English in the form of formal education in schools or other institutions, which is generally close to explicit learning in the Korean educational environment. Some of the participants, however, had experience of living in an English-speaking country or of immersion education, and there were some cases where they received

[11] Regular classes from other institutions mean studying to prepare for university entrance exams or English language proficiency tests.

elementary education in an English-speaking country under the age of 15. Also, some participants responded that they mainly learned English using media or books. These participants might be regarded as learning English through implicit learning. However, considering that their experience of living in English-speaking or immersion education was relatively short (less than a year) and all of them received formal education in Korea as well, they are likely to learn English in the form of explicit learning more than implicit learning.

Most of the participants reported that they did not use English in their daily lives except for reading (Table 5.1). The respondents who used English for less than an hour a week account for 63.3%, 80.0% and 80.0% of all responses in listening, writing, and speaking, respectively. This indicates that they rarely encounter English in a natural setting. However, in reading, about 63 percent of the participants answered that they used English for more than 1 hour per week, suggesting there is at least minimal exposure to English in reading.

(per week)	Listening	Reading	Writing	Speaking
1 hour or less	63.33%	36.67%	80.00%	80.00%
1 to 10 hours	26.67%	43.33%	13.33%	20.00%
10 to 20 hours	10.00%	20.00%	6.67%	0.00%
20 hours or more	* Three participants were removed from the analysis			

Table 5.1. Frequency of current use of English

5.5.2. Stimuli

88 pairs of words (Table 5.2) were created using the perceptual strength norms constructed in Experiment 1 and served as stimuli for Experiment 2. In Experiment 1, perceptual strengths of six senses —vision, hearing, smell, taste, touch, and interoception— were evaluated and compared across L1 English, L1

Korean and L2 English. From these multi-lingual perceptual strength norms, 44 auditory dominant words, 44 interoceptive dominant words, and 44 control words (mostly visual dominant words) were selected, based on perceptual strength, dominant modality, translatability, frequency, familiarity, and word length. The word selection process was as follows. First, to create the candidate list for auditory dominant words, the words with dominant modality (the modality with the highest perceptual strength level among the six senses) of audio or high audio strength and low interoception in both L1 Korean and L2 English perceptual strength norms were selected. Then, those with low translatability (below 6 out of 11) were removed from the list. Translatability was measured by how many 11 independent raters agreed on a Korean translation of the word. In the translation task conducted prior to Experiment 1 (see the method section of Experiment 1 for details), the raters translated 1000 English words into the first Korean words that came to their mind. So, if seven raters, for example, translated the word, *vision*, into "시각" in Korean, the translatability of *vision* is 7 out of 11. In a similar manner, the candidate for the interoceptive dominant words was chosen. These two candidate groups were heterogeneous in terms of audio and interoception perceptual strength (Table 5.3, $p<0.05$). After creating the candidate lists, 44 auditory words and 44 interoceptive words were selected from each candidate group, matched for frequency, familiarity, and word length. Word frequency for Korean and English were measured with the Python package WordFreq (Speer et al., 2018) and familiarity were evaluated by the participants in Experiment 1. There was no significant difference in Korean frequency, English frequency, Korean familiarity, English familiarity, Korean word length, and English word length between two finalist word groups ($p>0.05$). 44 control words were randomly selected from perceptual strength norms, excluding audio and interoception words, based on perceptual strength, translatability, frequency, familiarity, and word length. Control words had low interoception and audio strength, while translatability, frequency, familiarity, and word length were comparable for those of the audio and interoception group (*p>0.05*).

Audio	airport, argument, audio, blues, broadcast, command, complain, conference, criticism, dialogue, explain, feedback, interview, joke, lecture, loud, lyrics, media, movie, music, musical, news, noise, parade, police, praise, question, quiet, riot, shout, sing, speak, speech, sport, story, suggestion, swear, theater, thunder, train, video, voice, wave, word
Interoception	addiction, afraid, angry, ashamed, asleep, awkward, belief, bored, confused, depression, disease, emotional, exhausted, faith, fever, forget, glad, gravity, guilt, health, hunger, hurting, injured, lonely, nervous, nightmare, oxygen, pain, pride, pulse, rage, recover, recovery, relief, rest, shame, stomach, stress, tension, think, thought, tired, unhappy, worry
Control	ahead, android, background, behavior, bible, bitter, column, company, cotton, couch, dollar, eight, entrance, explore, fiction, five, historic, historical, japan, juice, mark, math, nine, number, optical, pocket, population, sealed, seek, senior, setting, soft, store, structure, sugar, summit, tape, technical, three, trick, university, visit, vote, wisdom

Table 5.2. Vocabulary list for Experiment 2

The 132 selected words were used to create two different prime-target pairs for the repetition and non-repetition conditions within each language. To balance between two pairs, first, the audio and interoception words are randomly divided into two groups (A and B). Then, the first half (group A) is assigned to the repetition condition of stimulus list 1 and the other half (group B) is assigned to the non-repetition condition of list 1; in list 2, B is assigned to repetition condition and A to non-repetition condition. Control words worked as prime words for non-repetition conditions. Each non-repetitive prime was matched for word length with target word as possible. This process resulted in two stimulus lists of 88 prime-target pairs for each language. A non-words list was also generated, which included 62 pseudo words created by reconstructing 1000 words used in Experiment 1. With 88 word-pairs and 62 non-word pairs, the total number of items in each stimulus list was 150. These lists were counterbalanced across participants.

		Audio	**Interoception**	**Control**	**t-stat†**
	Translatability	8.36	7.59	8.36	-1.93
	Concreteness	3.56	2.65	3.67	-5.41***
English	**Interoceptive**	1.07	3.34	0.96	22.25***
	Auditory	3.95	1.52	1.42	-22.77***
	Visual	3.24	2.60	3.58	-3.85***
	Frequency	4.70	4.54	4.72	-1.73
	Familiarity	2.84	2.82	2.79	-0.7
	Length	6.11	6.30	6.16	0.53
Korean	**Interoceptive**	0.88	3.17	0.71	21.94***
	Auditory	4.07	1.77	1.51	-19.23***
	Visual	3.43	3.08	3.87	-2.03*
	Frequency	4.57	4.39	4.54	-1.08
	Familiarity	2.66	2.64	2.64	0.12
	Length	2.91	2.88	2.86	-1.48

Table 5.3. T-statistics for vocabulary. Translatability, concreteness, interoceptive, auditory, and visual perceptual strength, frequency, familiarity, and length in English and Korean for audio, interoception and vision word list. † t-statistics for the t-test between audio and interception words. * $p<0.05$, ** $p<0.01$, *** $p<0.001$

5.5.3. Procedure

The experiment consisted of a pre-survey about language history, EEG experiment, and post-word test. A pre-survey was created based on Li et al. (2020) to examine the participants' L2 learning background, and frequency of daily

English usage, and English proficiency, including the CEFR self-assessment grid for proficiency evaluation. Participants performed it online prior to the EEG experiment or in the laboratory during EEG device setting. For the EEG experiment, the participants visited the laboratory, where they were given the instruction on the experiment and signed a consent form. Those who agreed to participate in the experiment were seated in an acoustically and electrically shielded room, where they wore EEG equipment. They were provided the instruction about a lexical decision task, in which they silently read the word presented on the screen and decide whether it is a word or non-word (below for the detail). In order to minimize EEG noise, participants were asked to remain still and not to blink their eyes if possible while the stimulus was being presented. Although they were allowed to take a break whenever they want, they were recommended to use four short breaks given during the experiment. They could selectively use the given break time, and most of them used it for a brief break. Before the main experiment, they first performed 10 practice trials to get accustomed to the lexical decision task. The main experiment was divided into two sessions for counterbalancing between languages. Half of the participants performed a L2 English task in the first session and a L1 Korean task in the second session, while the other participants did vice versa. A longer break was given between sessions. During this break, a noisy channel, if any, was adjusted and the participants were asked whether there were any issue or questions regarding the experiment. After they finished the EEG experiment, a post word test was administered while the device was removed. The purpose of the word test was to determine whether they knew the English words used in the experiment and what Korean they were translating those words into. They were asked to translate each English word into the first Korean that came to mind and were instructed to skip any words they did not know. All procedure was approved by and conducted in accordance with SNUIRB (the ethical committee of Seoul National University). Participants were paid 30000 won (about 25 dollars) for the participation.

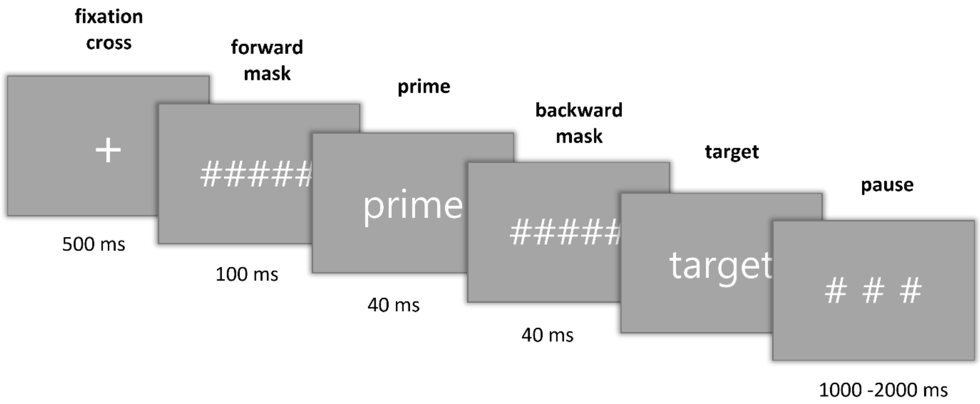


Figure 5.2. A sequence of individual trials of the lexical decision task.

The structure of each individual trial of lexical decision task was as follows (Figure 5.2). First, a fixation cross was presented in the center of the screen for 500 msec, in order to inform the participants of the position of the stimulus and induce their concentration. This was followed by a prime word lasting for 40 msec, which was covered by a forward mask of 100 msec and a backward mask of 40 msec. The masks consisted of seven hash marks, which were able to cover all Korean and English words of the material. Following the prime, the target word was shown, which continued to be present until the participant made a judgment about it. This response time was used as a dependent measurement for behavioral results. The participants were instructed to press the green button on the response pad if the target was a word, and the red button for a nonword. After their response, the target disappeared and three hash marks appeared. The duration of the hash marks ranged from 1000 msec to 2000 msec, and the average was 1500 msec.

5.6. Data analysis process

5.6.1. Data trimming and analysis for lexical decision tasks

The data for lexical decision task were trimmed as follows. First, responses with an unrealistically high reaction time, 10 seconds or more, were eliminated (0.03% of all data). Then, responses longer or shorter than average by 3 standard deviations were removed (1.85%). Finally, incorrect responses were excluded from the response time analysis, which accounted for 2.19% of all data. Inferential statistical analysis was performed using a mixed effect linear regression model with R statistical packages, lme4 (Bates et al., 2015) and lmerTest (Kunzetsova et al., 2017). The analysis was conducted in three steps. At the first stage, the differences between languages and between words and non-words were examined. *Language* and *word type* were included as fixed effects, and *participants* and *items* were considered as random effects. In the second step, with the data excluding non-words, whether there was a difference in repetition and modality between languages and how repetition effect was different depending on modality was determined. *Language*, *Modality* (audio vs. interoception) and *Repetition* were fixed effects, while random effects were *articipants* and *items*. As a final stage, the same analysis as in the second step was conducted within each language. Fixed effects were *Modality* and *Repetition*, whereas random effects were *participants* and *items*.

5.6.2. EEG recording

The participants' EEG was recorded from 64 active scalp electrodes embedded in an ActiCap Slim (BrainProducts GmbH, Gilching, Germany). The electrodes were arranged following the international 10-20 system with electrode FCz as the online reference. In order to detect noise caused by eye movement and blinking, four additional electrodes were attached to the outer canthus of both eyes (HEOG: Horizontal Electro-oculogram) and under and above the right eye (VEOG: Vertical Electro-oculogram). All electrode impedances were set below 10 kOhms prior to and during recording. The EEG signal was amplified through actiCHamp (BrainProducts GmbH), digitized at a sampling rate of 500 Hz, and recorded through BrainVision Recorder (BrainProducts GmbH).

5.6.3. Pre-processing

EEG data were pre-processed with MNE-Python package (Gramfort et al., 2013). The pre-processing was conducted as follows. First, raw data were digitally filtered with a zero-phase Finite Impulse Response (FIR) filter (high-pass: 0.1 Hz, 12 dB/octave; low pass (for ERP): 30 Hz, 12 dB/octave; low pass (for time-frequency analysis and source estimation): 100 Hz, 12 dB/octave). Then, ocular-artifacts were semi-automatically removed by independent component analysis (ICA) with picard algorithm (Ablin, Cardoso & Gramfort, 2018), which is newer, supposed to be faster than other methods such as FastICA and Infomax, and more robust in cases where the sources are not completely independent. Corrected EEG data were re-referenced to the average reference. For ERP and source estimation analysis, continuous EEGs were segmented in the range of -200 to 800 msec from the onset when the target word was presented, and the baseline was corrected in the range of -200 msec to 0 msec. For Time-frequency analysis, they were segmented between -500 and 1500 msec, since longer segments are required compared to ERP in order to analyze the low frequency band. After segmentation, for the baseline correction, the mean of baseline values was subtracted from trials and divided by the mean of baseline values (‘percent’ mode of baseline correction from MNE package). The time window of -500 to -180 msec from the onset was used as the baseline because this was where a black screen was presented before masked priming so that the participants' neural activity was minimized. Automated artifact rejection technique was applied to segmented data by autoreject python package (Jas et al., 2017), which automatically interpolates artefactual channels and deletes contaminated epochs with Bayesian optimization method. Also, incorrect responses in the lexical decision task were eliminated from the trials. If more than 12.5 % (average loss percentage) of the epochs of the data were removed as incorrect answers and through artifact rejection, the whole data of that participant was excluded from the analysis (7 participants). After artifact rejection, for the ERP and source estimation analysis, ERPs were extracted by averaging remaining segments (of trials with correct answers) independently in each participant for each condition. For the time-frequency analysis, each pre-processed trials were transformed into the power spectrograms, and then averaged by conditions and participants (See *5.6.5. Time-frequency analysis procedure* for detail).

5.6.4. Statistical procedure for ERP analysis

To statistically confirm the difference in the ERP responses among the conditions, repeated measures analyses of variance (ANOVA) was performed on the target word, using R (R Core Team, 2019). When the sphericity assumption was violated, the Greenhouse-Geisser correction was applied. Two separate ANOVAs were conducted each for midline and lateral electrode sites (Figure 5.3). The analysis for lateral sites included factors of *Modality* (Audio *vs*. Interoception), *Repetition* (repetition *vs*. nonrepetition), *Region* (anterior, central, and posterior), and *Hemisphere* (left *vs*. right). The midline analysis included the same factors except *Hemisphere*. In addition, in order to examine the difference in Modality by the repetition effect in more detail, the dataset was divided into the repetition-only and the non-repetition-only data, and a separate analysis was performed on the modality within each data. These separate analyses included *Modality*, *Region*, and *Hemisphere* (for lateral sites) as factors. All the analyses were conducted in three time-windows: 100–250, 250–400, and 400–550 msec, which are known to be associated with the repetition priming effect, modality specific neural responses and early and late language processing, based on previous studies (Eddy et al., 2006; Holcomb, & Grainger, 2007; Trumpp et al., 2014; also see *5.4.2.1. ERP components* for detail).

Aside from parametric statistical analysis in specific scalp areas and time windows, cluster-based non-parametric permutation tests were performed across all electrode sites and time points with MNE-Python package (Gramfort et al., 2013). Cluster-based non-parametric permutation tests can effectively solve the multiple comparisons problem especially when the spatiotemporal sources of the effect are unknown in advance (Maris & Ooostenveld, 2007). This additional statistical analysis will verify the robustness of the results with a data-driven statistical test (See detail for *3. Neurological Background*). Permutation tests were applied to the difference waves of *Repetition* (repetition *vs.* non-repetition), *Repetition* within each modality (e.g. audio repetition *vs.* audio non-repetition), *Modality* (audio *vs.* interoception) and *Modality* within each repetition (e.g. audio repetition *vs.* interoception repetition) in each language.

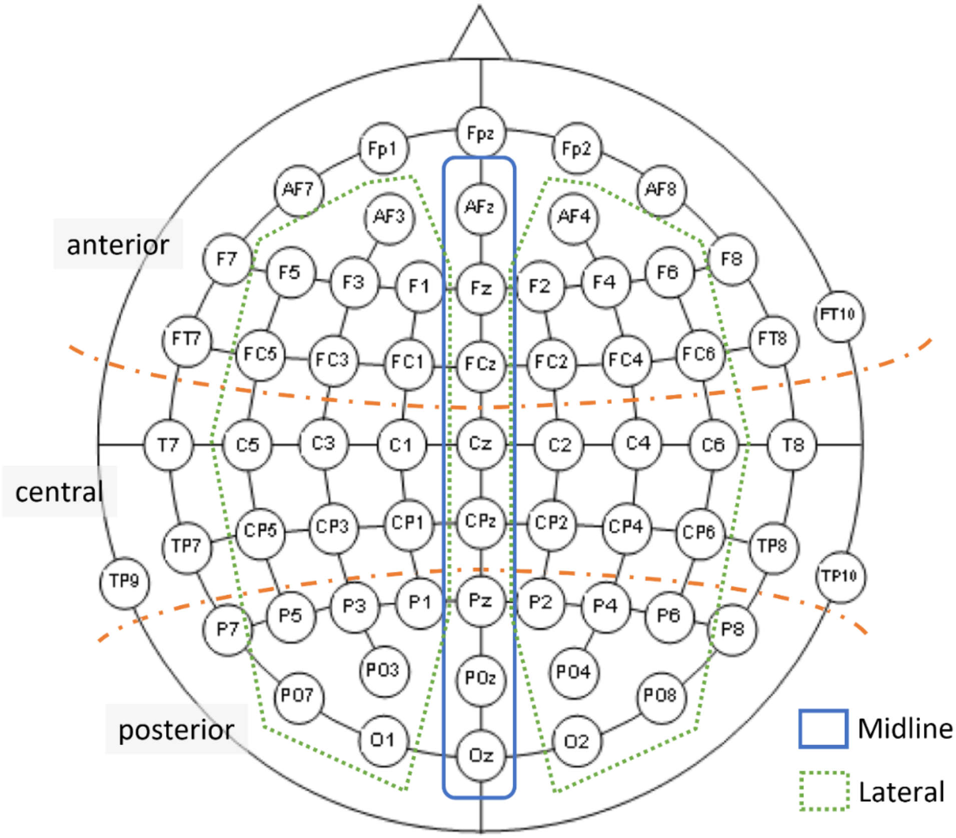


Figure 5.3. Electrode montage and two analysis sites (Midline / Lateral) used for ANOVAs.

5.6.5. Time-frequency analysis procedure

For time-frequency analysis, each preprocessed trial was transformed into power spectrograms. For the transformation, Morlet wavelets (width of 3–4 cycles, linearly increased according to the frequency) were used for the low frequency range (4–35 Hz), while multitaper approach (width of 4–7 cycles) for the high frequency range (25–70 Hz). A Morlet wavelet is a type of window function employed to preserve temporal localization which is lost in the Fourier transform process (Cohen, 2014). It has the appearance of a sine wave in the middle but smoothly tapers out to zero at both ends. As a result, while preventing the loss of temporal information, this compensates for the drawback of employing a sine wave as a window function that both ends of the window are shaped with sharp edges. The multitaper approach is to generate multiple power spectra by applying multiple tapers and use their averaged spectra as the output (Cohen, 2014). Since this technique is useful for enhancing the signal-to-noise ratio of the frequency representation, it helps analyze noisy trials such as higher frequency activities. However, this transformation is ineffective at lower frequencies below 30 Hz because the averaging several tapers might undermines frequency isolation. Thus, in this work, Morlet wavelet is employed for the lower frequency band, while the multitaper method is used for the higher frequency. After transforming each trial into power spectra, they were averaged by conditions and participants, resulting in averaged power spectrograms for each condition and participant. These average power spectrograms were used as dependent variables for the statistical analysis. For the statistical analysis, cluster-based non-parametric permutation tests were conducted across all frequencies and time points for selected electrodes (Frontal: F3, Fz, F4; Central: C3, Cz, C4; Parietal: P3, Pz, P4; Occipital: O1, Oz, O2) with MNE-Python package (Gramfort et al., 2013). The tests were performed between *Modality* (audio *vs.* interoception) in each language.

5.6.6. Source estimation procedure

To estimate the neural sources responsible for the difference in the ERP and time-frequency analysis across modalities and languages, the distributed source analysis was conducted with MNE-Python package (Gramfort et al., 2013). Sources were estimated from the grand-averaged ERP for each modality (audio and interoception) and the difference between the repetition and non-repetition condition within each modality (audio repetition *vs* audio non-repetition and interoception repetition *vs* interoception non-repetition). The time window was designated as 140-160 msec from onset in Korean and 160-180 msec in English, where the global field power (GFP) was highest in each language. GFP is a measure of signal agreement over the whole scalp. If the signals of the sensors vary, the GFP at that point will be larger, which is known to reflect some brain activities that require further examination (Gramfort et al., 2013; Lehmann & Skrandies, 1984). Based on ERP and time-frequency results, two frequency bands of interest where the difference between audio and interoception was found were selected for the analysis (lower frequency band: 0.1-30 Hz, gamma: 30-70 Hz). For the computation of the forward model, a typical adult template MRI or "fsaverage" was used for the head model (Fischl, 2012). For the inverse modelling, the time window of -500 to -180 msec from the onset (black screen before the masked prime is present) was used as the pre-stimulus baseline to estimate the noise regularization parameters. In order to increase the reliability of source estimation, several inverse modelling methods available in the MNE-Python package such as dSPM, MNE, eLORETA, and sLORETA were used and the results were cross-validated. The obtained strength of the source estimates of each frequency band was then averaged. Source spaces with strength greater than 95% of the data percentiles were considered as sources for the conditions. Only when the four methods consistently estimated similar regions as the sources, they were regarded as a reliable outcome and reported.

5.7. Results

5.7.1. Response time and accuracy

The accuracy of the lexical decision task was 97.91% ($SE=14.32\%$, $min=92.33\%$, $max=100\%$) on average (Table 5.4). The average rate of correct answer was significantly higher in Korean than in English ($estimate=1.51$, $SE=0.48$, $t=3.17$, $p<0.01$) and in word condition than in non-words regardless of languages (estimate=2.63, SE=0.48, t=5.53, $p<0.001$). However, there was no significant difference between *Modality*, *Repetition*, or *Language* in the analysis within only word condition ($p>0.05$).

	Total	**Word**					**Nonword**
			Audio		Interoception		
			Rep.	Non-Rep.	Rep.	Non-Rep.	
Korean	98.57% (0.12)	99.33% (0.08)	99.38% (0.30)	99.17% (0.37)	99.59% (0.26)	99.17% (0.30)	97.47% (0.37)
English	97.23% (0.16)	98.59% (0.12)	97.71% (0.60)	98.95% (0.42)	99.17% (0.54)	98.54% (0.52)	95.13% (0.52)

Table 5.4. Mean and standard deviation of the accuracy rate for a lexical decision task. Rep.: Repetition, Non-Rep.: Non-Repetition.

The average of the overall reaction time was 765.53 msec ($SE=3.50$) (Table 5.5). For response time, the main effect of *Language* ($estimate=-187.61$, $SE=10.06$, $t=-18.66$, $p<0.001$) and *Word type* ($estimate=-206.06$, $SE=10.06$, $t=-$

20.49, p<0.001), as well as the significant interaction between *Language* and *Word type* (*estimate=148.06, SE=20.11, t=7.36, p<0.001*), were found. The response time in Korean was significantly faster than in English and to words than to non-words. This difference between word and non-word were greater in English.

The second analysis which excluded non-word conditions revealed a main effect of language (*estimate=-114.0, SE=12.21, t=-9.33, p<0.001*) where the participants responded faster in Korean than in English within the word condition, and a main effect of repetition (*estimate=-62.47, SE=7.47, t=-8.36, p<0.001*), which suggested that the repetition effect occurred both in Korean and English. There was no significant interaction between Modality and Repetition (*p<0.1*).

	Total	**Word**					**Nonword**
			Audio		Interoception		
			Repetition	Non-Rep.	Repetition	Non-Rep.	
Korean	682.59	630.16	591.43	670.40	598.72	660.30	759.54
	(5.56)	(5.64)	(9.26)	(10.19)	(8.95)	(8.63)	(6.31)
English	842.42	759.54	696.47	765.03	732.17	772.44	1002.64
	(3.87)	(4.67)	(10.30)	(12.14)	(11.38)	(10.98)	(10.08)

Table 5.5. Mean and standard deviation of the response times for lexical decision tasks. Non-Rep.: Non-Repetition.

5.7.2. ERPs

5.7.2.1. Korean

Repetition

In the time window of 100-250 msec, there was a main effect of Repetition (Midline: *F(1, 22) = 4.89, p<0.05*) and an interaction between Repetition and Region (Midline: *F(1.59, 35.03) = 8.66, p<0.01*; Lateral: *F(1.5,*

$33.1) = 9.12$, $p<0.01$) (Figure 5.4, 5.5, 5.7). Repetition conditions induced positivity compared to non-repetition conditions over the anterior (Midline: $F(1, 22) = 9.03$, $p<0.01$; Lateral: $F(1, 22) = 9.56$, $p<0.01$) and central regions (Midline: $F(1, 22) = 7.08$, $p<0.05$; Lateral: $p>0.05$), while negativity over posterior regions (Midline: $F(1, 22) = 7.80$, $p<0.05$; Lateral: $F(1, 22) = 11.3$, $p<0.01$). The waveform at this time window peaking around at 150 msec following the first negative peak (N1) resembles a typical P2 component, with a larger P2 in the repetition condition, consistent with previous research (Misra & Holcomb, 2003).

In the time window of 250-400 msec, which corresponds to N400, a main effect of Repetition (Lateral: $F(1, 22) = 7.17$, $p<0.05$) and an interaction between Repetition and Region (Midline: $F(1.48, 32.55) = 10.5$, $p<0.001$; Lateral: $F(1.51, 32.55) = 10.5$, $p<0.001$) was found. However, different from the previous time window, neural responses were more negative in the repetition condition over anterior regions (Midline: $F(1, 22) = 14.1$, $p<0.001$; Lateral: $F(1, 22) = 8.47$, $p<0.01$) but positive over central (Midline: $F(1, 22) = 5.33$, $p<0.05$; Lateral: $p>0.05$) and posterior regions (Midline: $F(1, 22) = 8.25$, $p<0.01$; Lateral: $F(1, 22) = 11.2$, $p<0.01$). In this time-window, the repetition condition shows smaller N400 (less negative) compared to non-repetition conditions over the central and posterior regions, consistent with previous studies (Holcomb & Grainger, 2007; Misra & Holcomb, 2003).

In the time window of 400-550 msec, the results revealed an interaction between Repetition and Region both in midline ($F(2, 44) = 9.96$, $p<0.001$) and lateral analysis ($F(2, 44) = 10.1$, $p<0.001$). There was significant negativity in repetition conditions over anterior region (Midline: $F(1, 22) = 21.9$, $p<0.001$; Lateral: $F(1, 22) = 13.9$, $p<0.001$), positivity over central (Midline: $p>0.05$;

Lateral: $p>0.01$) and posterior region (Midline: $p>0.05$; Lateral: $F(1, 22) = 11.4$, $p<0.01$). This time-window coincides with part of N400 and LPC. In the early part of this time window, the N400-like response of the previous time-window sustains, but in the latter part (LPC), the difference in amplitude by repetition effect decreases. In the masked priming paradigm, LPC is less likely to be modulated by repetition, so the result is compatible with previous studies (Misra & Holcomb, 2003).

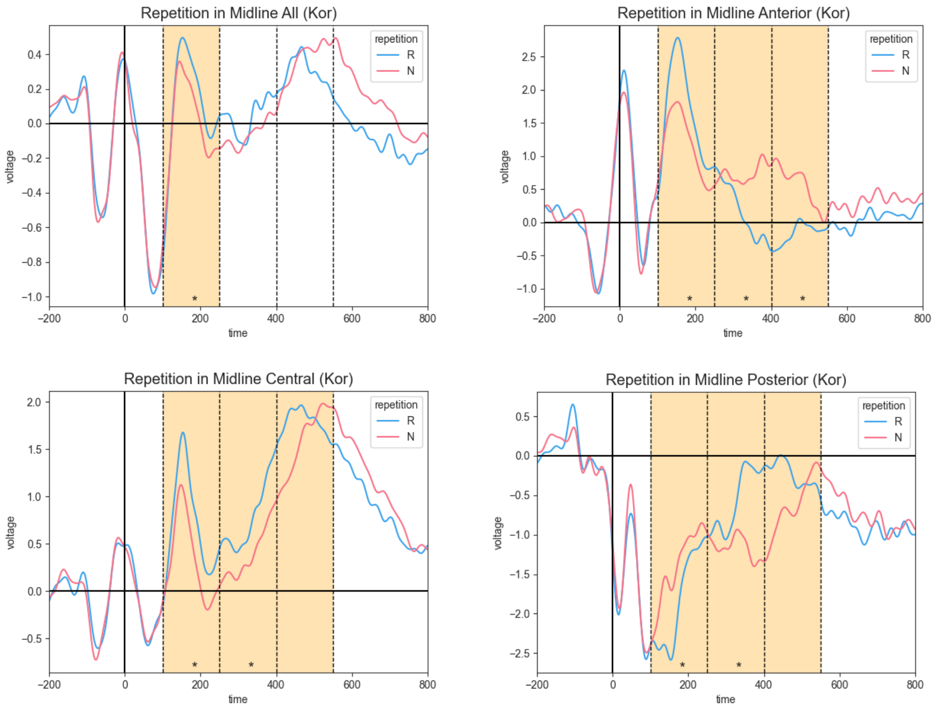


Figure 5.4. Grand-averaged ERPs from midline as a function of Repetition (repetition: blue line *vs.* non-repetition: red line) and Region in Korean. Orange-shaded areas indicate the time windows where a main effect or significant interaction was found. An asterisk represents a significant difference found in post-hoc analysis. R: Repetition, N: Non-repetition.

To summarize, in all three-time windows, the repetition effect was found within both modalities. This effect, in earlier time window, induced positivity in anterior and central area and negativity in posterior area, while in later time window, negativity over the anterior regions and positivity over central and posterior regions. The timing, polarity, and topography of repetition effect seem to indicate larger P2 and smaller N400 of repetition conditions.

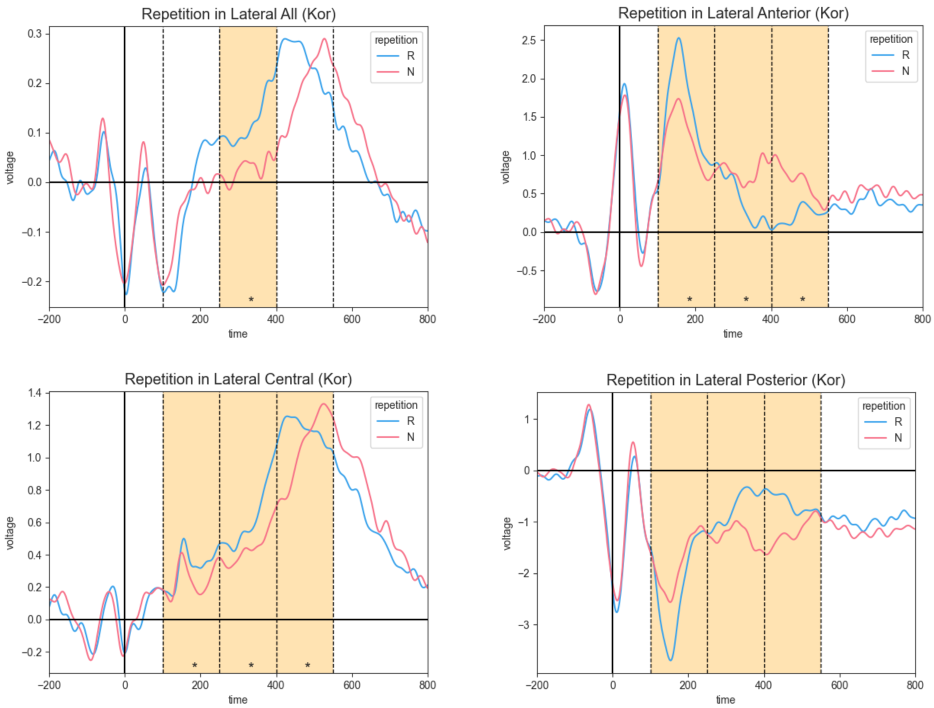


Figure 5.5. Grand-averaged ERPs from lateral electrode sites as a function of Repetition (repetition: blue line *vs*. non-repetition: red line) and Region in Korean. R: Repetition, N: Non-repetition.

Modality

There was a main effect of Modality in all three time-windows at midline sites (100-250 msec: $F(1, 22) = 4.34$, $p<0.05$; 250-400 msec: $F(1, 22) = 4.52$, $p<0.05$; 400-550 msec: $F(1, 22) = 8.79$, $p<0.01$) (Figure 5.6, 5.7). Audio conditions showed more positivity than interoception from early to late time windows. Summarizing the result based on ERP components, it is larger P2 for audio (more positive), larger N400 for interoception (more negative), and larger LPC for audio (more positive).

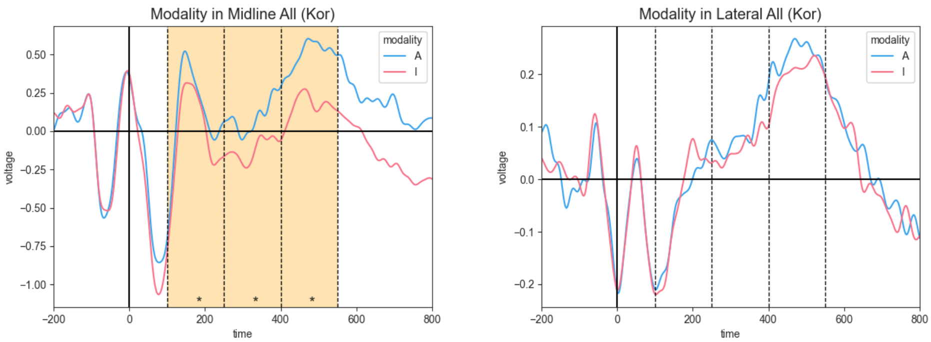


Figure 5.6. Grand-averaged ERPs from midline and lateral electrode sites as a function of Modality (Audio: blue line *vs.* Interoception: red line) in Korean. Orange-shaded areas with an asterisk represent the time windows where a main effect was found. A: Audio, I: Interoception.

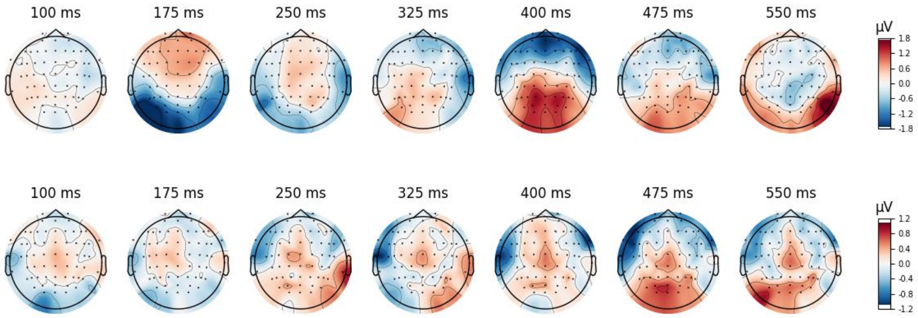


Figure 5.7. Topographic maps of a difference wave of [Top] Repetition (repetition minus non-repetition) and [Bottom] Modality (audio minus interoception) in Korean.

Modality x *Repetition*

The results at the lateral sites showed an interaction between Modality, Repetition and Hemisphere ($F(1, 22) = 4.47, p<0.05$) in the time window of 100-250 msec (Figure 5.8, 5.9). At the left hemisphere region, positive (marginal) repetition effect ($F(1, 22) = 3.22, p<0.1$) was found within audio conditions, whereas at the right hemisphere, more negative (but only numerically different) effect ($p>0.05$) was revealed. This finding indicates that repetition effect varied depending on Modality and Hemisphere, only present at the left hemisphere in audio conditions, but the significant effect was not found at the post-hoc analysis

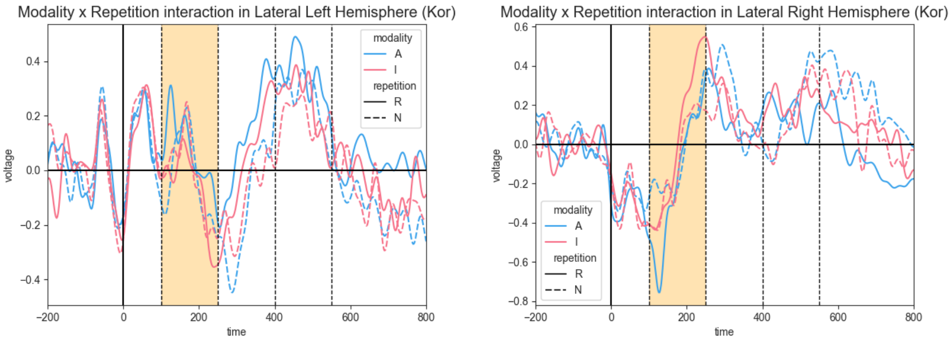


Figure 5.8. Grand-averaged ERPs from lateral electrode sites as a function of Modality (Audio: blue line *vs.* Interoception: red line), Repetition (Repetition: Solid line *vs.* Non-repetition: dotted line) and Hemisphere in Korean. Orange-shaded areas indicate the time windows where an interaction was found. A: Audio, I: Interoception, R: Repetition, N: Non-repetition.

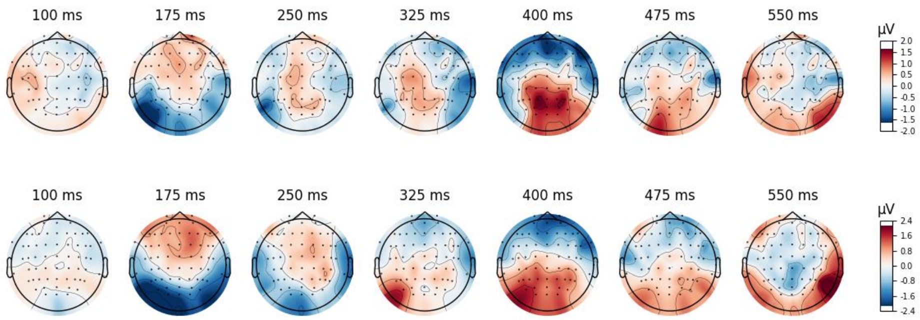


Figure 5.9. [Top] Topographic maps of a difference wave of audio repetition *vs.* non-repetition and [Bottom] interoception repetition *vs.* non-repetition in Korean.

Modality within repetition conditions and within non-repetition conditions

When differences between modalities were analyzed only with repetition conditions (Figure 5.10, 5.12), a significant interaction between Modality and Hemisphere ($F(1, 22) = 4.53, p<0.05$) was found at the lateral sites in the time window of 100-250 msec. This may be a result of the opposite polarity in the two hemispheres, a larger positivity for audio in left hemisphere and for interoception in right hemisphere, but neither was significant. In the time-windows of 250–400 msec ($F(1, 22) = 9.39$, $p<0.01$) and 400–550 msec ($F(1, 22) = 9.81$, $p<0.01$), there was a main effect of modality at midlines sites, which represented larger negativity for interoception (N400) and larger positivity for audio (LPC), respectively. In summary, within the repetition conditions, audio and interoception exhibit a distinct response from the early time windows, and the pattern was similar to the modality differences within all repetition conditions. On the other hand, no significant effect was found between the two modalities within non-repetition conditions ($p>0.1$) (Figure 5.11, 5.12).

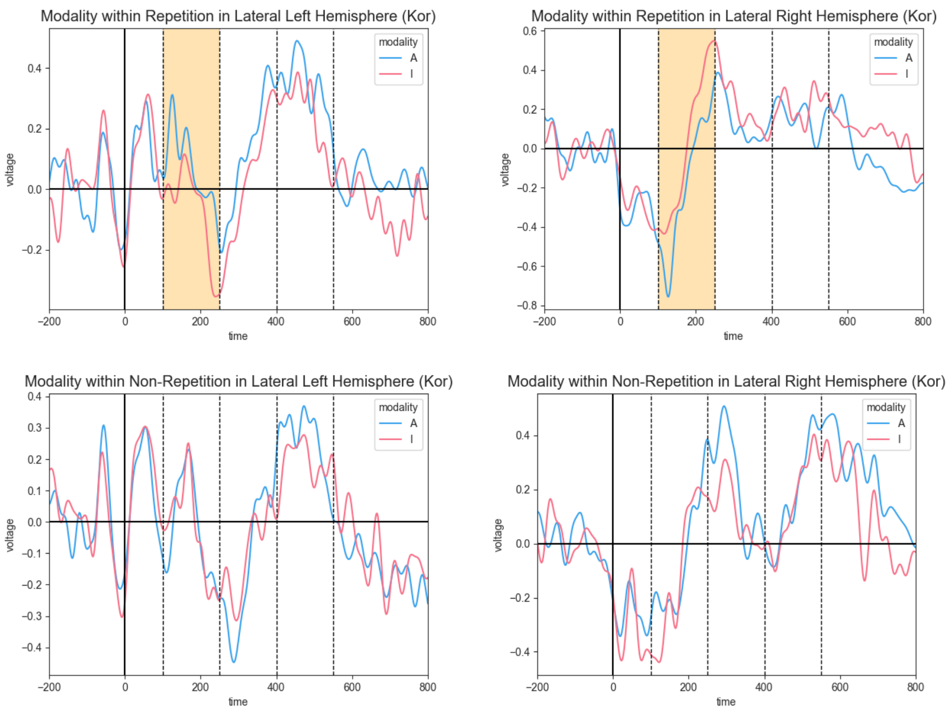


Figure 5.10. Grand-averaged ERPs from lateral electrode sites as a function of Modality (Audio: blue line *vs.* Interoception: red line) and Hemisphere [Top] only within the repetition conditions in Korean. Orange-shaded areas indicate the time windows where an interaction was found. [Bottom] Only within the non-repetition conditions in Korean. A: Audio, I: Interoception.

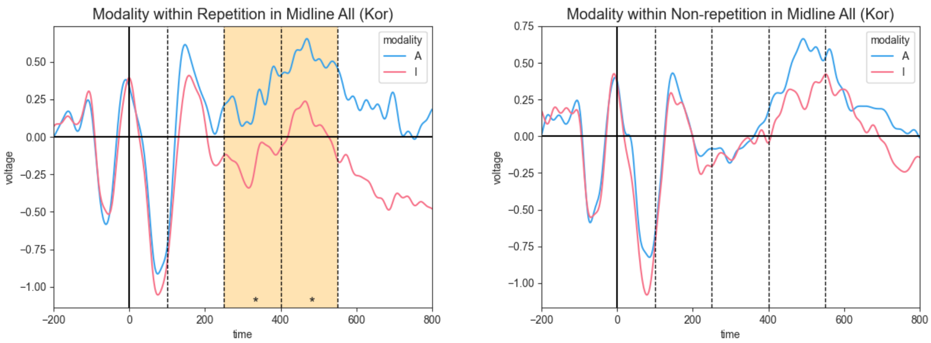


Figure 5.11. Grand-averaged ERPs from midline electrode sites as a function of Modality (Audio: blue line *vs.* Interoception: red line) only within the repetition conditions (Left) and only within the non-repetition conditions (Right) in Korean. Orange-shaded areas with an asterisk represent the time windows where a main effect was found. A: Audio, I: Interoception.

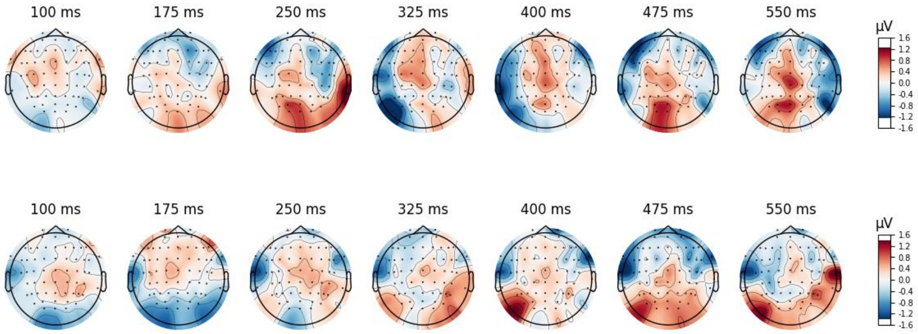
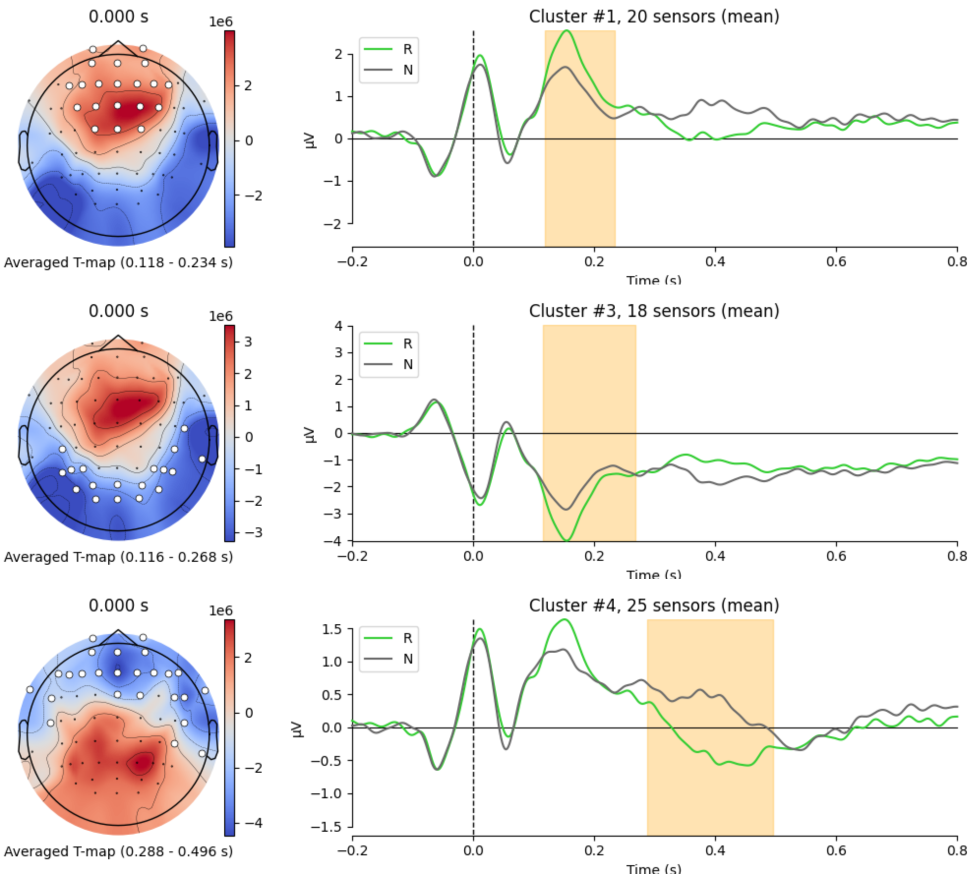


Figure 5.12. [Top] Topographic maps of a difference wave of audio *vs.* interoception within the repetition conditions and [Bottom] within the non-repetition conditions in Korean.

Non-parametric cluster Permutation test

Cluster permutation test revealed four significant clusters between repetition and non-repetition conditions (Figure 5.13). Cluster #1 and #3 in the early time window occupy a largely mutually exclusive domain, showing bipolar activities. Cluster #1 (Figure 5.13-A) contains 22 sensors distributed with positivity over the frontal and central regions, persisting from 118 msec to 234 msec, while

Cluster #3 (Figure 5.13-B) is distributed over posterior regions with negativity. The polarity pattern of voltage is typical of average-referenced data, and it is highly likely to reflect the activity of the same source generator. These two clusters seem to be related to the repetition effect in in early time windows in ERP analysis. In the later time window, Cluster #4 and #2 (Figure 5.13-C, D) showed another bipolar pattern of neural activities. Compared to the early time window, the difference is that the fronto-temporal area exhibits negativity while the central-posterior region has positive activity. This pattern is also consistent with the results of ERP analysis.



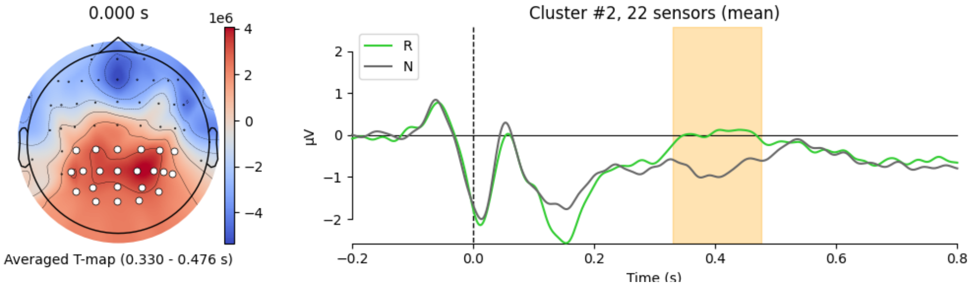
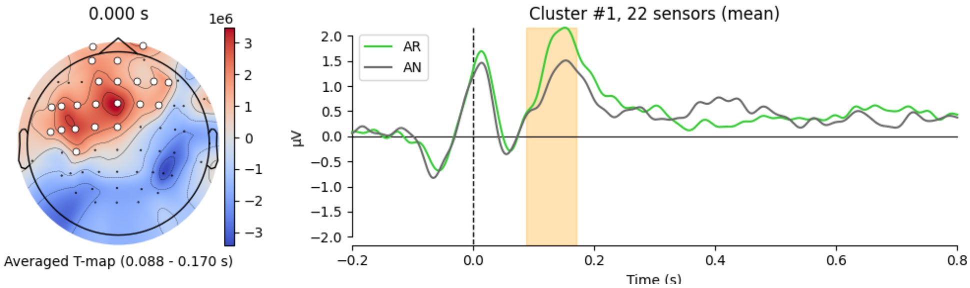


Figure 5.13. A-D. (From top to bottom). Results of the cluster permutation tests for the repetition (green line) vs. non-repetition (gray line). Four figures are sorted according to the time window and distribution, not the cluster number. The number of the clusters does not imply order or intensity but randomly set. The left side of each figure represents topographic map of clusters. White circles are electrodes included in the cluster. The right of each figure is ERP waveforms of repetition and non-repetition conditions. A colored shading indicates the significant time span for the cluster.

The permutation test within each modality also shows similar patterns of the clusters between repetition and non-repetition conditions. Three significant clusters were found between audio repetition and non-repetition conditions (Figure 5.14). Cluster #1 of audio (Figure 5.14-A) demonstrates the similar repetition effect in the early time window, but began slightly earlier; the distribution was more concentrated in the left hemisphere; and the bipolar activity did not form a statistically significant cluster. Cluster #3 and #2 (Figure 5.14-B, C) of audio reflect bipolar activities in the late time window—positivity over frontal regions and negativity over central and posterior area.



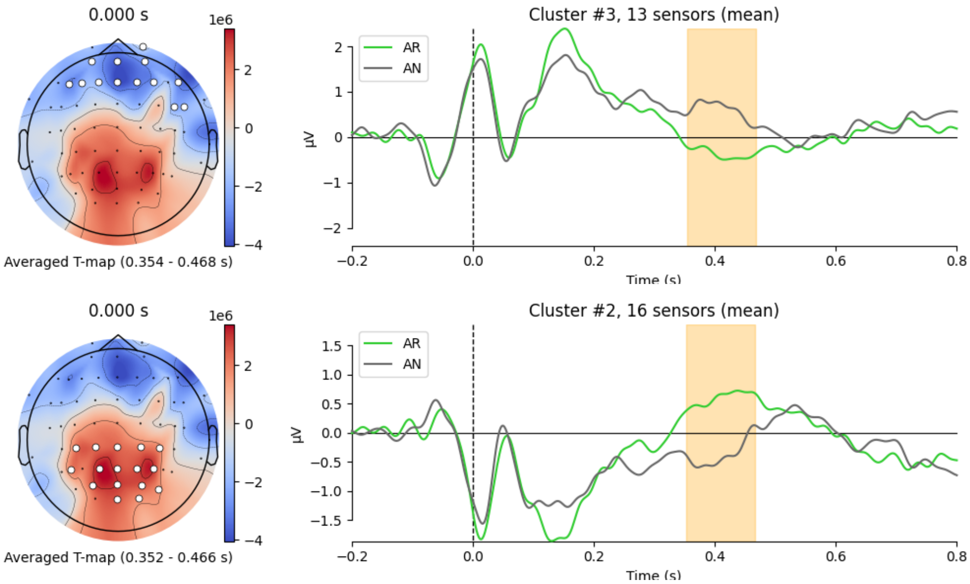


Figure 5.14. A-C. Results of the cluster permutation tests for the repetition (green line) vs. non-repetition (gray line) within audio modality.

There were four significant clusters for interoception (Figure 5.15). Cluster #1 and #3 of interoception (Figure 5.15-A, B) revealed the polarity pattern in the early time window. Compared to the clusters of audio in the early time window, Cluster #1 of interoception was dispersed less in the left hemisphere, beginning 40 msec late, and a significant cluster was also found in the posterior region, reflecting bipolar activities. Clusters #4 and #2 (Figure 5.15-C, D) of interoception indicate bipolar activity in the late time window, with positivity over the frontal areas and negativity over the central and posterior regions, which was similar with the clusters of audio in late time windows.

On the other hand, no significant cluster was seen between audio and interoception, inconsistent with the ANOVA results, which revealed a main effect of Modality.

To summarize, the results of the non-parametric cluster permutation test revealed the repetition effect of each modality, which is generally similar to P2 and

N400 effect identified in ANOVA in terms of polarity, timing, and topography. However, no modality effect was found.

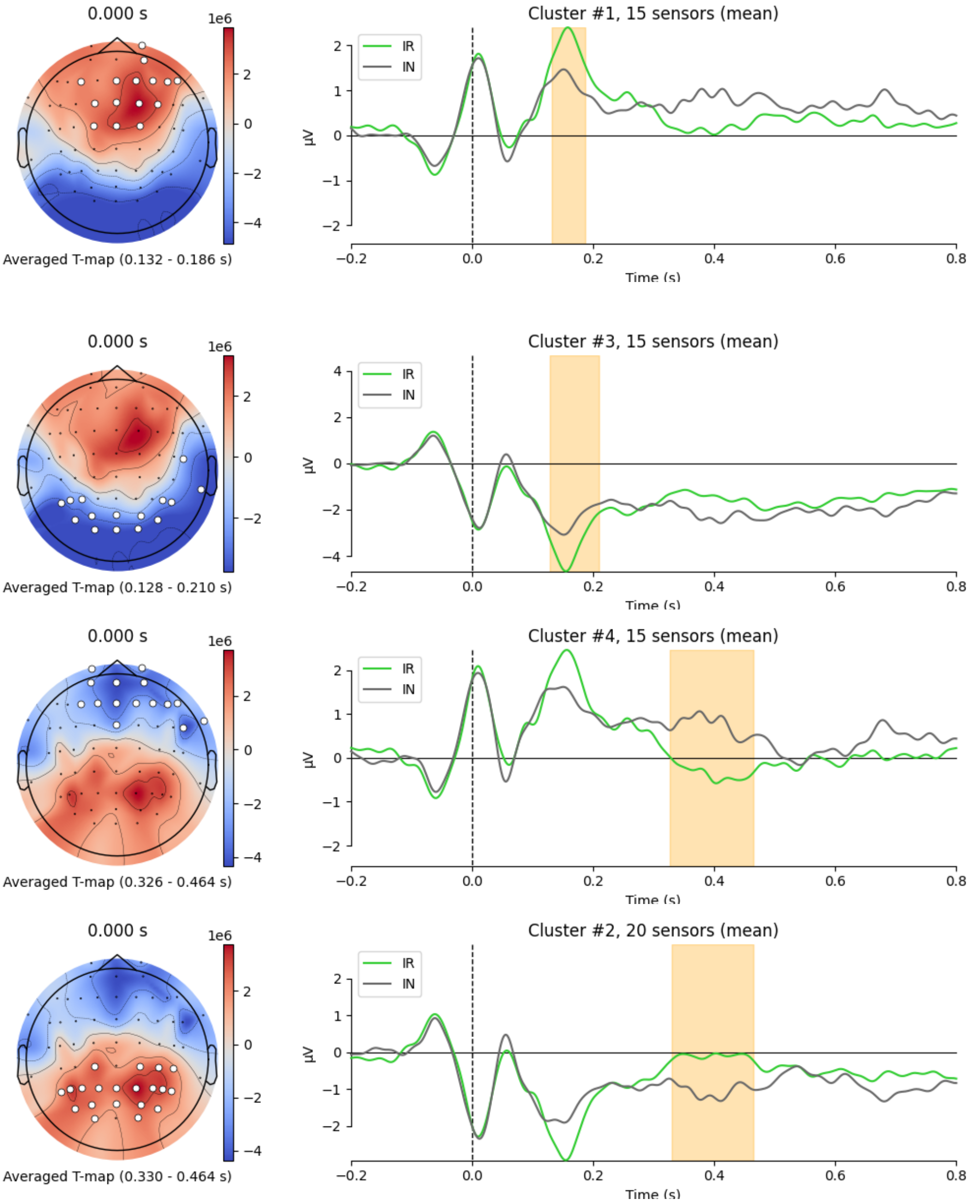


Figure 5.15 A-D. Results of the cluster permutation tests for the repetition (green line) vs. non-repetition (gray line) within interoception modality.

5.7.2.2. English

Repetition

For midline analysis, a main effect of Repetition was found in all three time windows (100-250 msec: $F(1, 22) = 4.67, p<0.05$; 250-400 msec: $F(1, 22) = 6.10, p<0.05$; 400-550 msec: $F(1, 22) = 8.74, p<0.01$) (Figure 5.16, 18). Repetition conditions elicited positive neural activities compared to non-repetition ones. In the time window of 100-250 msec, there was a significant interaction between Repetition and Hemispheres ($F(1, 22) = 5.73, p<0.001$). In the left hemisphere, repetition condition showed negativity, but positivity in the right hemisphere. However, both effects were not significant ($p>0.1$). In the time windows of 250-400 msec and 400-550 msec, a significant interaction between Repetition and Region was revealed in the midline electrode sites (250-400 msec: $F(2, 44) = 4.00, p<0.05$; 400-550 msec: $F(1.41, 30.95) = 4.53, p<0.05$). In the both windows, repetition conditions induced significant positivity over central (250-400 msec: $F(1, 22) = 5.01, p<0.05$; 400-550 msec: $F(1, 22) = 11.2, p<0.01$) and posterior regions (250-400 msec: $F(1, 22) = 9.45, p<0.01$; 400-550 msec: $F(1, 22) = 7.89, p<0.01$), while negativity over anterior areas ($p<0.05$).

In summary, the results showed that larger positivity (P2) for repetition condition in the time window of 100-250 msec and smaller negativity (N400) for repetition condition in the time window of 250-400 and 400-550 msec over central and posterior regions. This is consistent with the findings from Korean data and previous studies.

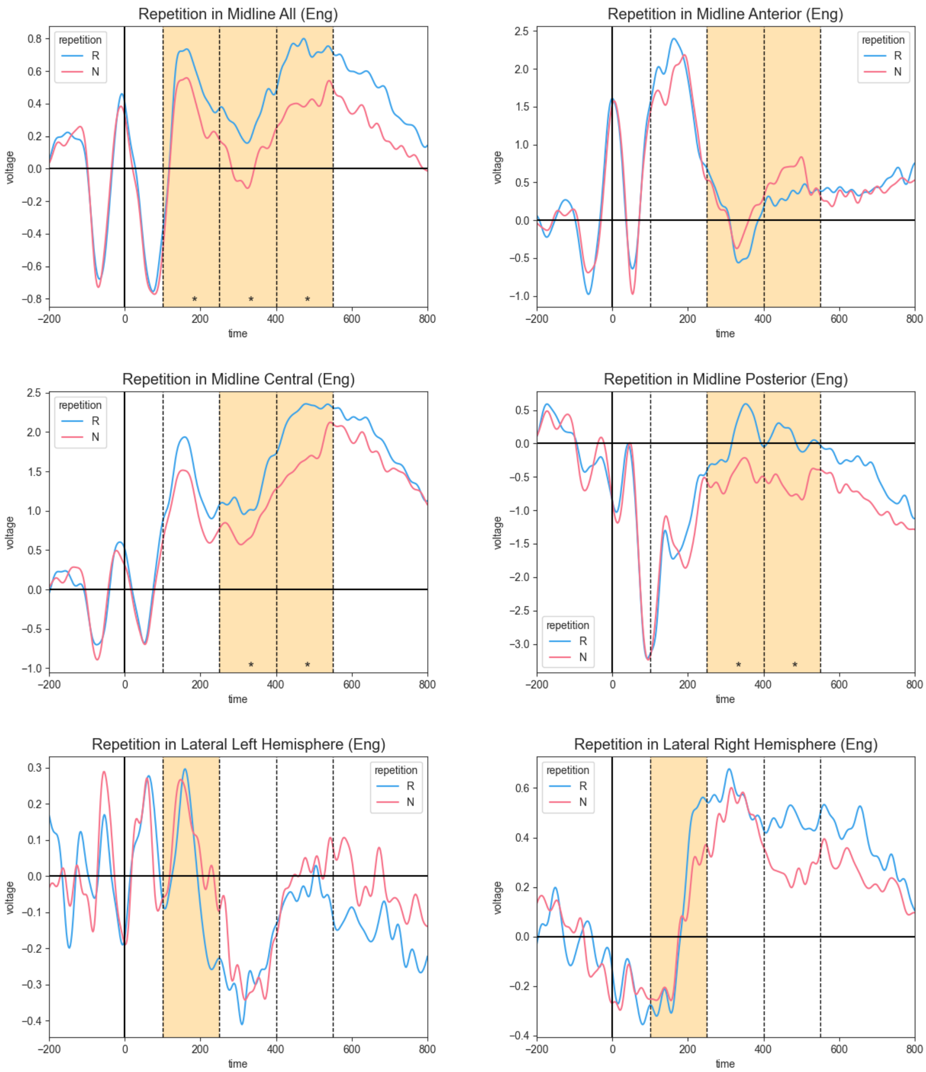


Figure 5.16. Grand-averaged ERPs from midline and lateral electrode sites as a function of Repetition (repetition: blue line *vs.* non-repetition: red line) and Region in English. R: Repetition, N: Non-repetition.

Modality

Similar to the findings in Korean, Audio was significantly more positive than interoception in all three time-windows for midline analysis (100-250 msec: $F(1, 22) = 6.25, p<0.05$; 250-400 msec: $F(1, 22) = 9.93, p<0.01$; 400-550 msec: $F(1, 22) = 5.52, p<0.05$) (Figure 5.17, 18). Besides a main effect of modality, an interaction between Modality and Hemisphere was found in the time window of

100-250 msec (*F(1, 22) = 4.36, p<0.05*) and 250-400 msec (*F(1, 22) = 6.15, p<0.05*). In both time window, there was significant difference between audio and interoception only in the right hemisphere (100-250 msec: *F(1, 22) = 4.48, p<0.05*; 250-400 msec: *F(1, 22) = 6.42, p<0.01*), where interoception elicited more negative neural responses. These results show larger P2 and LPC for audio (more positive) and larger N400 (more negative) for interoception similarly in Korean.

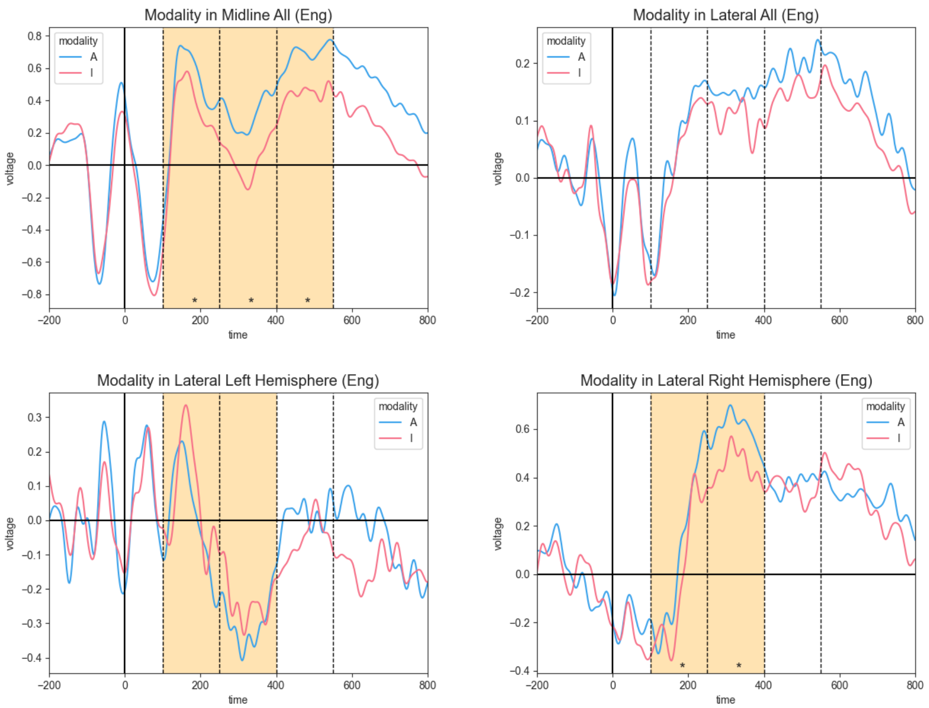


Figure 5.17. Grand-averaged ERPs from midline and lateral electrode sites as a function of Modality (Audio: blue line *vs.* Interoception: red line) and Hemisphere in English. Orange-shaded areas with an asterisk represent the time windows where a main effect was found. A: Audio, I: Interoception.

Modality x *Repetition*

In the first time window, there was an interaction of Repetition, Modality, and Region for midline electrode sites (*F(1.22, 26.93)=4.67, p<0.05*) (Figure 5.19, 21). The separate analysis over the anterior regions showed a significant interaction between Modality and Repetition (*F(1, 22)=4.61, p<0.05*), which is accounted for by positivity of repetition condition only within interoception. On the other hand, over the posterior region, a marginal interaction between Modality and Repetition (*p<0.1*) was found, which is related with positive repetition effect of audio condition. In the time window of 250-400 msec, the analysis for the lateral electrode sites revealed a significant interaction between Modality and Repetition (*F(1, 22)=4.78, p<0.05*) (Figure 5.20, 21). This seemed to be explained by the repetition effect with opposite polarities, negativity for audio and positivity for interoception, but the difference was not statistically confirmed. In summary, first, a difference in repetition effect between audio and interoception was found in earlier time windows at the midline sites. Over the anterior regions, there was a repetition effect in interoception, while in the posterior, it was present in the audio. In addition, lateral analysis revealed contrasting repetition effects between audio and interoception in the second time window.

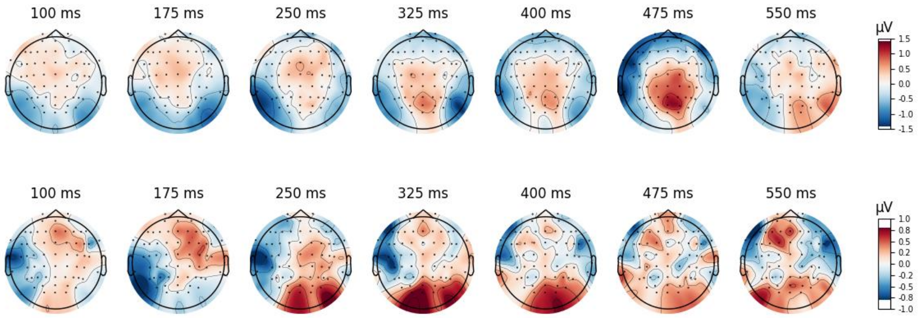


Figure 5.18. [Top] Topographic maps of a difference wave of Repetition (repetition *minus* non-repetition and [Bottom] Modality (audio *minus* interoception) in English.

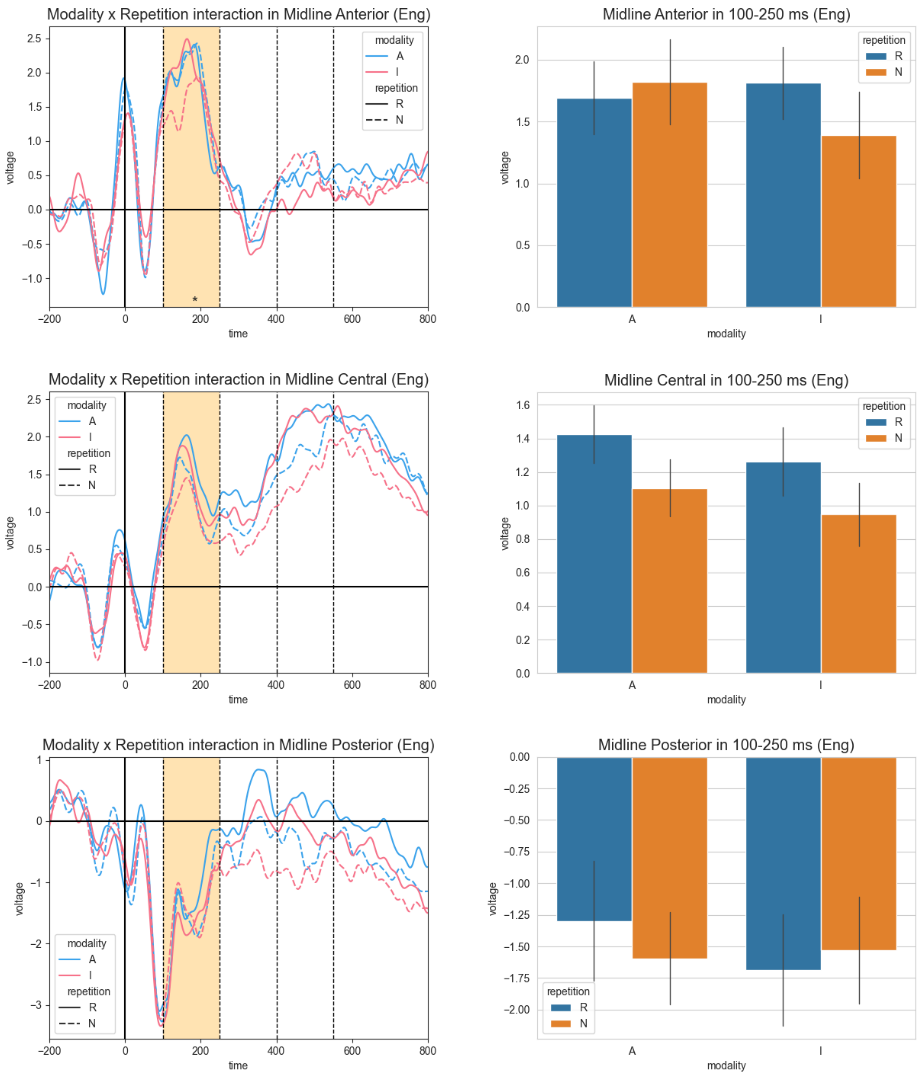


Figure 5.19. [Left] Grand-averaged ERPs from midline electrode sites by Region (anterior, central, posterior) as a function of Modality (Audio: blue line *vs.* Interoception: red line), Repetition (Repetition: Solid line *vs.* Non-repetition: dotted line) [Right] Mean amplitude across Modality and Repetition by Region within the time window of 100-250 msec. A: Audio, I: Interoception, R: Repetition, N: Non-repetition.

Modality within repetition conditions and within non-repetition conditions

In the time window of 250-400 msec, there were significant differences between modalities in both repetition and non-repetition conditions (Figure 5.22, 5.23). In non-repetition conditions, a main effect of Modality (Midline: (*F(1,* 22)=6.91, p<0.05*); Lateral: *F(1, 22)=4.56, p<0.05*) and a significant interaction between Modality and Hemisphere (*F(1, 22)=5.57, p<0.05*) were found. This indicated that the negativity of interoception (N400) was greater than that of audio, especially in the midline sites and right hemisphere. In the repetition conditions, a significantly larger negativity was found in interoception at the midline (*F(1, 22)=5.12, p<0.05*).

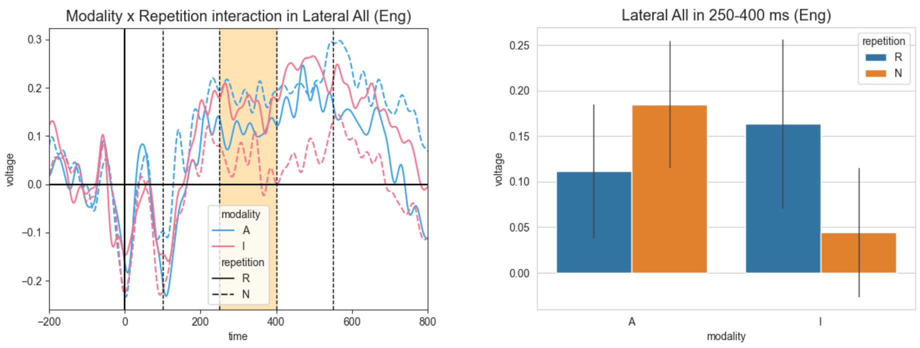


Figure 5.20. [Left] Grand-averaged ERPs from lateral electrode sites by Hemisphere (left, right) as a function of Modality (Audio: blue line *vs.* Interoception: red line), Repetition (Repetition: Solid line *vs.* Non-repetition: dotted line) [Right] Mean amplitude across Modality and Repetition by Region within the time window of 100-250 msec. A: Audio, I: Interoception, R: Repetition, N: Non-repetition.

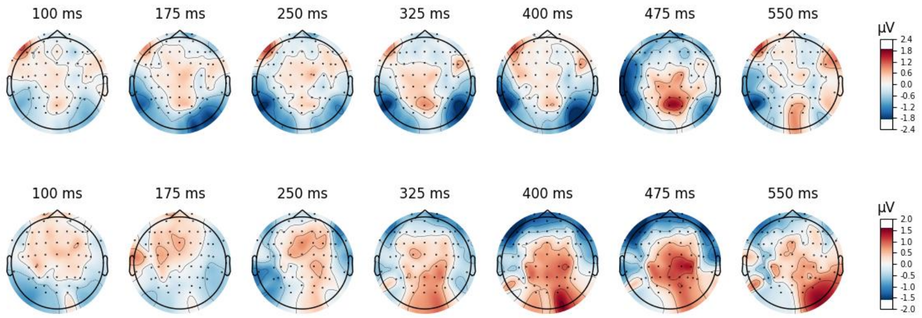


Figure 5.21. [Top] Topographic maps of a difference wave of audio repetition *vs.* non-repetition and [Bottom] interoception repetition *vs.* non-repetition in English.

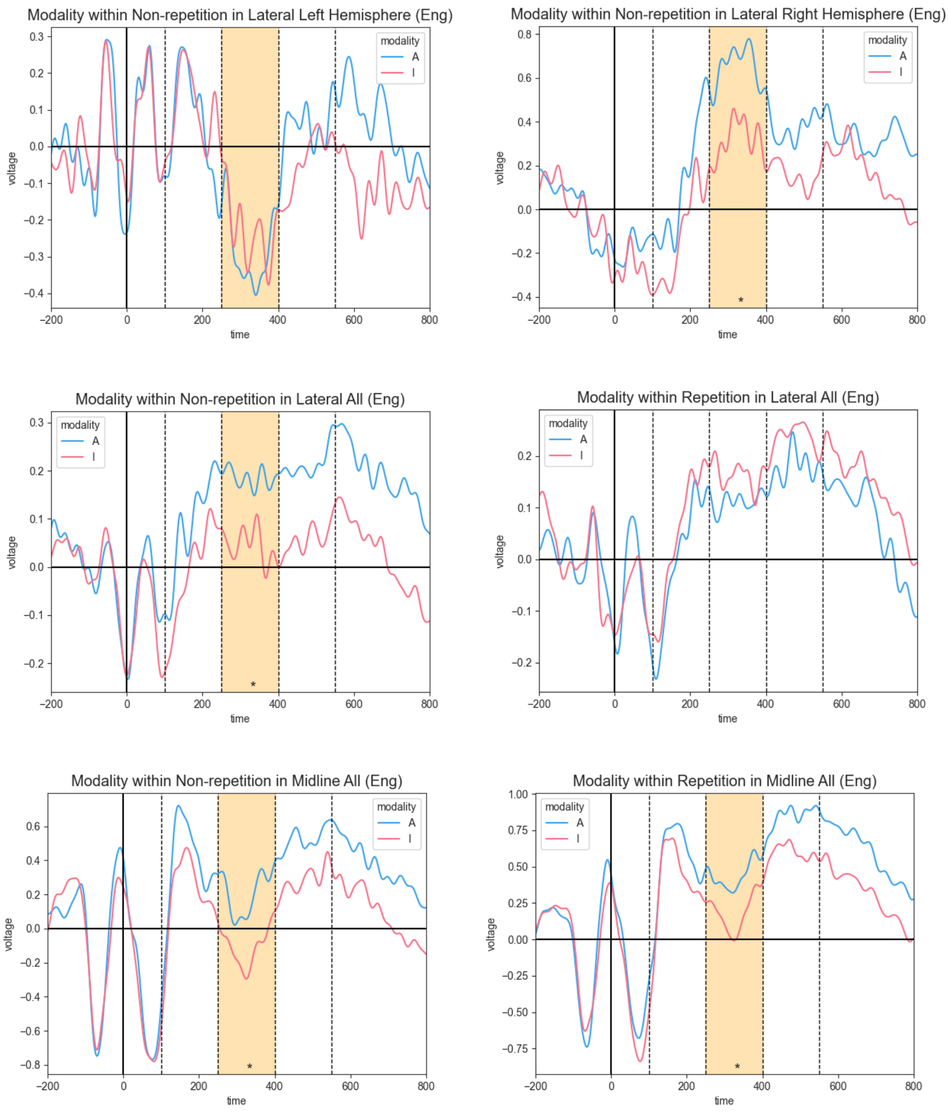


Figure 5.22. [Top] Grand-averaged ERPs in English from lateral electrode sites as a function of Modality (Audio: blue line *vs.* Interoception: red line) and Hemisphere only within the non-repetition conditions for left hemisphere (Left) and right hemisphere (Right). Orange-shaded areas with an asterisk indicate the time windows where a significant interaction was found. [Middle] for both hemispheres within the non-repetition conditions (Left) and within the repetition conditions (Right). [Bottom] Grand-averaged ERPs from midline electrode sites as a function of Modality (Audio: blue line *vs.* Interoception: red line) only within the non-repetition conditions (Left) and only within the repetition conditions (Right). A: Audio, I: Interoception.

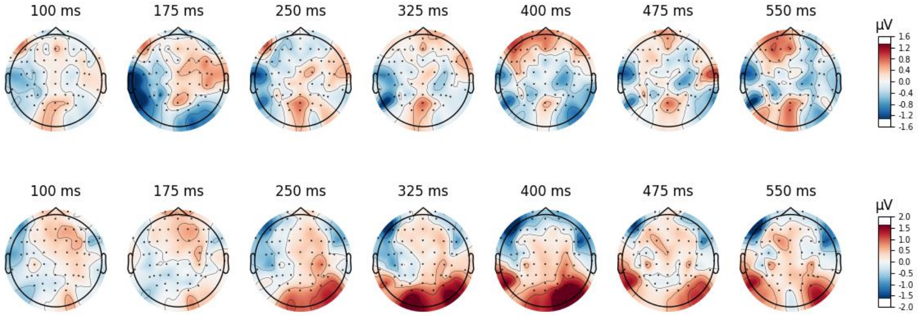
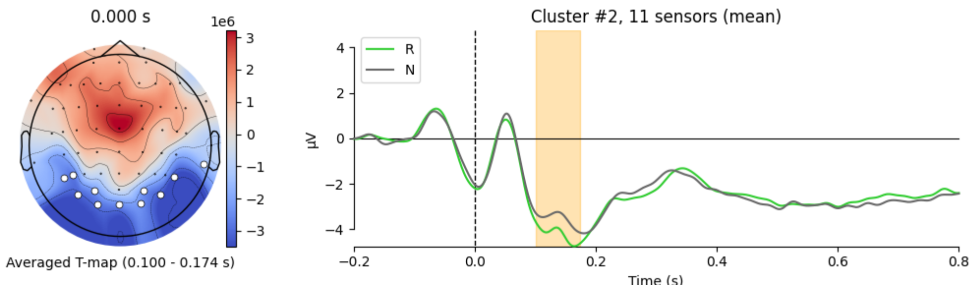


Figure 5.23. [Top] Topographic maps of a difference wave of audio *vs.* interoception within the repetition conditions and [Bottom] within the non-repetition conditions in English.

Non-parametric cluster permutation test

For English data, three significant clusters were found between repetition and non-repetition conditions. Cluster #2 (Figure 5.24-A) consists of 11 sensors spread negatively throughout the posterior areas, lasting between 100 and 174 msec. In later time windows, Cluster #1 and #3 (Figure 5.24-B, C) have a generally bipolar distribution. Cluster #1 demonstrated a negative distribution over the central and parietal regions, whereas Cluster #3 displayed a negative distribution over the frontal and parietal areas slightly centered on the left hemisphere. Overall, the results of the permutation test for English data show similar timing and distribution to those of the Korean language, and seem to agree with the findings of the ANOVA test.



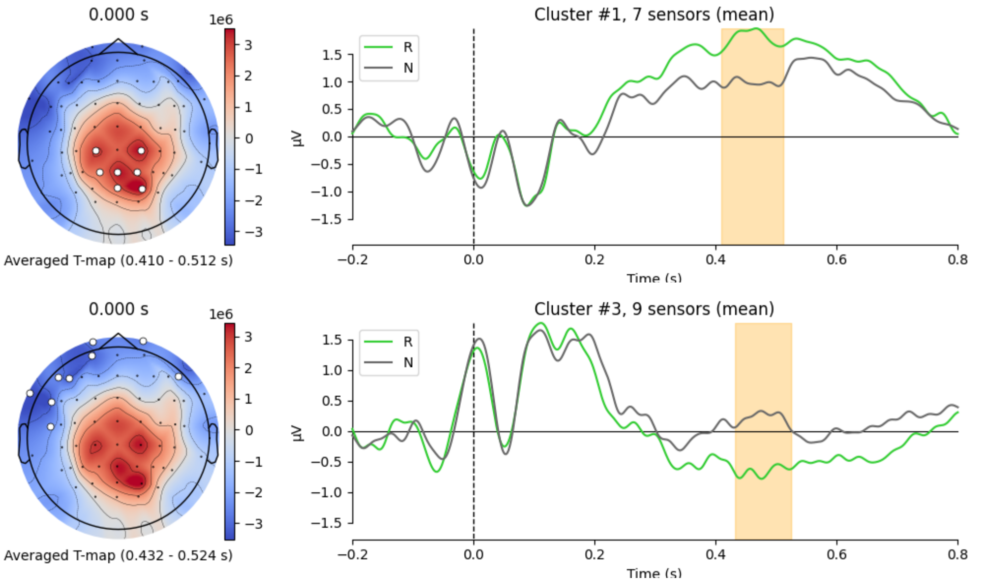


Figure 5.24 A-C (From top to bottom). Results of the cluster permutation tests for the repetition (green line) vs. non-repetition (gray line) in English data.

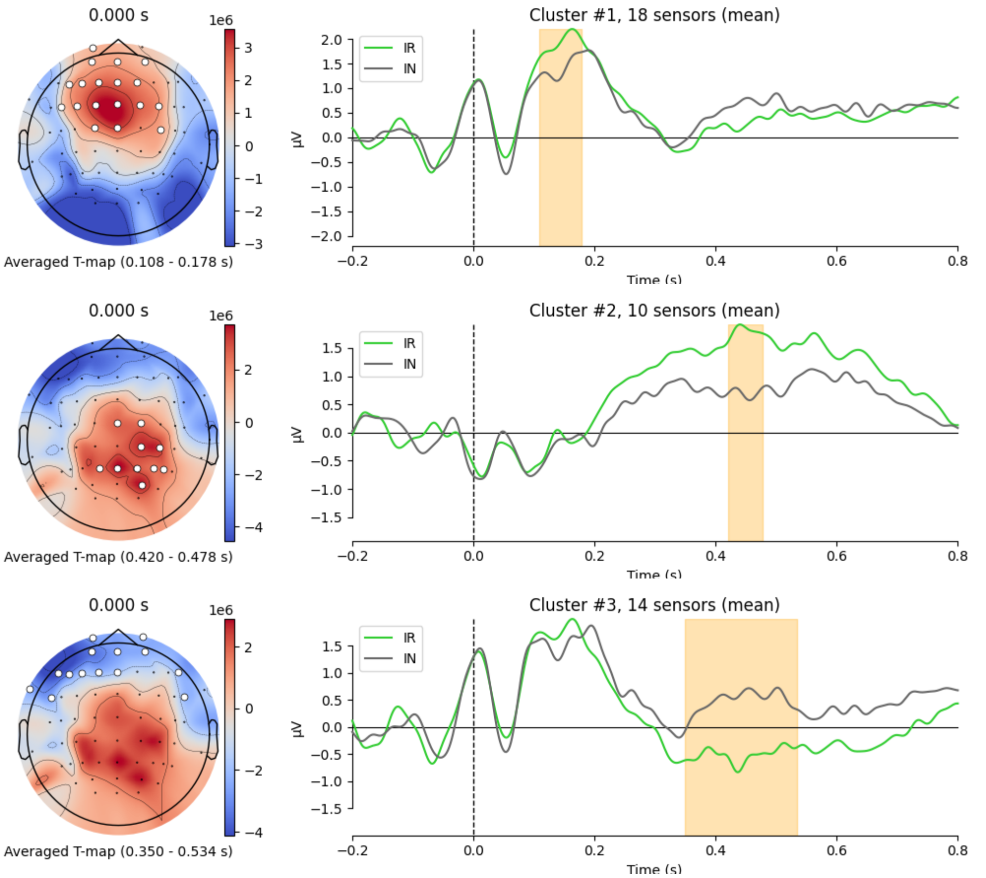


Figure 5.25. A-C. Results of the cluster permutation tests for the repetition (green

line) vs. non-repetition (gray line) within interoception modality.

When the repetition and the non-repetition condition were compared within each modality, three significant clusters were found only in interoception. In the early time window, positivity of the fronto-central area was clustered (Cluster #1; Figure 5.25-A), and in the late time window, bipolar activities of positivity over frontal regions and negativity over central and posterior area were observed (Cluster #2 and #3; Figure 5.25-B, C). Overall, the distribution of the repetition effect compared within interoception is not noticeably different from when both modalities were included. However, no significant cluster was identified within the audio condition. This difference may suggest a greater repetition effect within interoception, but it was not statistically supported. This result is to some extent consistent with the ANOVA results where an interaction between Modality and Repetition was found, implying a greater repetition effect in interoception, but it is also not statistically supported.

5.7.3. Time-frequency analysis

5.7.3.1. Korean

The results of cluster-based non-parametric permutation test for time-frequency analysis showed a significant cluster at the occipital sites in the comparison of audio and interoception (Figure 5.26). The cluster is located within the frequency range of 35-64 Hz, which overlaps the (lower) gamma band (30–70 Hz), in the time window of 284-364 msec. No significant cluster was found between other conditions.

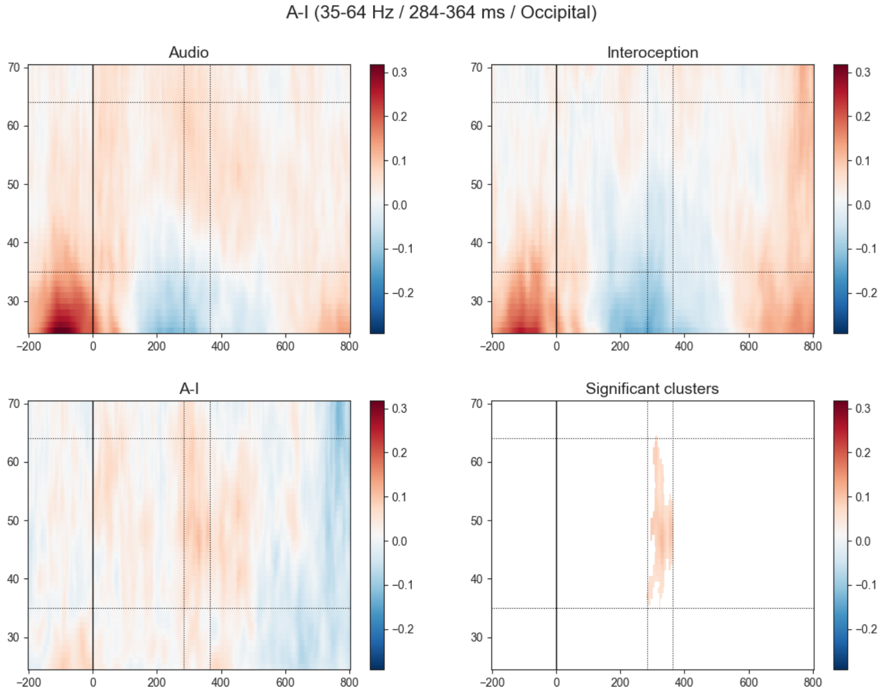


Figure 5.26. The results of the time-frequency analysis in Korean. [Top] An example of power spectrograms at the occipital sites (O1, Oz, O2). From the top-left, Audio, Interoception, Audio *minus* Interoception, and significant clusters of Audio *minus* Interoception. Significant clusters are formed in the frequency range of 35-64 Hz (gamma) and the time window of 284–364 msec. [Bottom] Topographical maps of the gamma band range for audio and interoception in the time window of significant clusters. A: Audio, I: Interoception

5.7.3.2. English

In the comparison of audio and interoception, a significant cluster was found at the parietal area (Figure 5.27) in the time window of 300-516 msec, within the range of 25–42 Hz from higher frequency analysis (multitaper) and in the time window of 352-420 msec, within the range of 21–34 Hz from lower frequency analysis (Morlet). This frequency range spans the beta (15–30 Hz) and gamma bands (30–70 Hz).

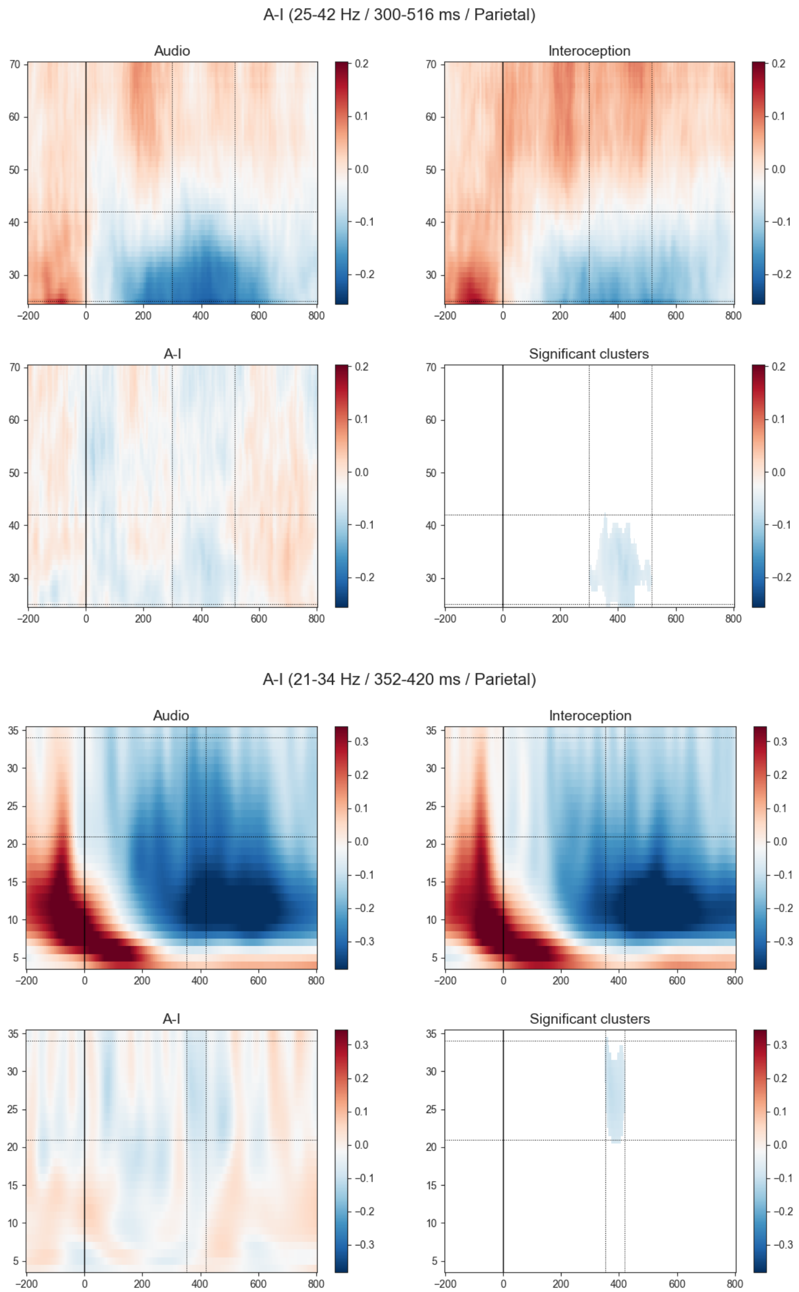


Figure 5.27. [Top] The results of the time-frequency analysis in English. An example of power spectrograms at the parietal site. From the top-left, Audio, Interoception, difference between Audio and Interoception, and significant clusters of Audio *minus* Interoception. A significant cluster is formed in the frequency range of 25-42 Hz (beta-gamma) and the time window of 300–516 msec. [Bottom] The results of the time-frequency analysis in English. A significant cluster in the frequency range of 21-34 Hz (beta) and in the 352–420 msec. A: Audio, I: Interoception.

5.7.4. Source estimation

5.7.4.1. Korean

For the analysis of gamma band (30-70 Hz), the sources for the audio condition (Figure 5.29) were estimated to be parts of the left and right temporal cortex (close to BA 21, 22, 41, 42; Figure 5.26 for Brodmann Areas) and right prefrontal cortex (BA 46) consistently across all four approaches. BA 41 and 42 are parts of the auditory cortex, while BA 21 and 22 is the auditory association cortex (Kraemer et al., 2005). The posterior part of BA 22 is where Wernicke's area reside, but it is generally located in the left hemisphere (Rasmussen & Milner, 1977). In this analysis, BA 21 and 22 were estimated as the sources only in the right hemisphere. BA 46 roughly corresponds with dorsolateral prefrontal cortex (DLPFC), which is known to be related with executive functions such as attention and working memory (Curtis & D'Esposito, 2003).

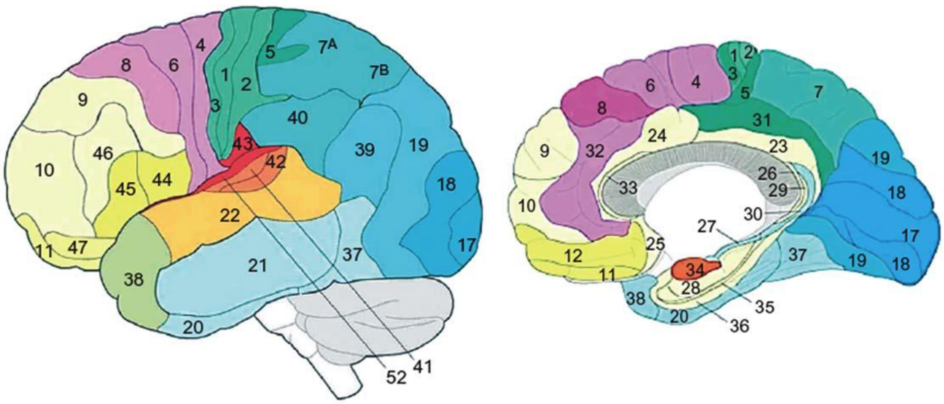


Figure 5.28. Brodmann areas. [Left] Lateral view of left hemisphere. [Right] Medial view of right hemisphere. Adapted from Gage & Baars (2018). (identical with Figure 5.1.)

The sources for interoception (Figure 5.29) were computed as parts of right temporal pole (BA 38), right prefrontal cortex (BA 11, 12, 45, 46, 47) and left prefrontal cortex (BA 10, 11). The temporal pole, or BA 38, is involed in many high-order cognitive functions such as visual processing, semantic processing in all modalities, and socio-emotional processing (Herlin, Navarro & Dupont, 2021) and regared as semantic hub (Patterson & Ralph, 2016). BA 11, 12, 45 and 47 are parts of inferior frontal gyrus and medial orbito prefrontal cortex. Inferior frontal gyrus is the part that overlies the insular and related with interoception (García-Cordero et al., 2017), whereas medial orbito prefrontal cortex is assumed to be related with taste, smell, emotion, social behavior and decision making regarding reward (Rolls, 2004; Rushworth et al., 2007).

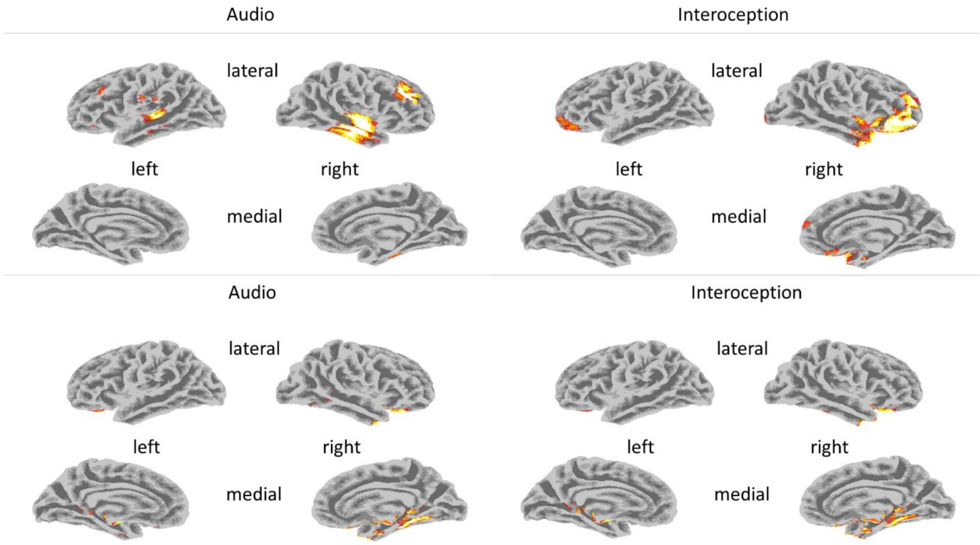


Figure 5.29. [Top] Examples of results of source estimation within gamma band (30-70 Hz) in Korean by eLORETA for the audio (Left) and interoception (Right). [Bottom] Source estimation within low frequency band (0.1-30Hz) for the audio (Left) and interoception (Right).

For the analysis of lower frequency band (0.1–30 Hz), the sources for both modalities (Figure 5.29) was commonly estimated to be parts of the right temporal pole (BA38) and medial temporal lobe (BA 20, 28, 34, 37). The sources in the medial temporal lobe are distributed over the regions encompassing the entorhinal cortex, fusiform gyrus, and parahippocampal gyrus. They are asuumed to memory input/retrieval and (word) recognition (Luck et al., 2010; McCandliss et al., 2003; Tsao et al., 2018).

In summary, audio and interoception were estimated to have different sources within the gamma band range, and some of the estimated sources seemed to be related to their respective modality (auditory cortex and inferior frontal gyrus). At lower frequencies, the two modalities similarly estimated the medial temporal lobe as the source.

In order to investigate whether the repetition effect is modulated by modality, the source of each modality was estimated within each repetition condition (Figure 5.30). Within the gamma band, the source of the audio repetition condition was estimated to be part of the left and right temporal lobe and right frontal lobe. These are related to audio cortex, comaprable to the estimation of the entire audio condition. On the other hand, in the non-repetition condition, a part of the pre-frontal cortex was presumed to be the source. In interoception, parts of the right temporal pole and the right prefrontal cortex, including the inferior frontal gyrus and medial orbito prefrontal cortex, were estimated as sources in both repetition and non-repetition conditions. One difference was that in the repetition condition, the estimates around the temporal pole were larger and stronger compared to the non-repetition conditions. Source estimates at lower frequencies are the same as those for each modality, so the report is omitted.

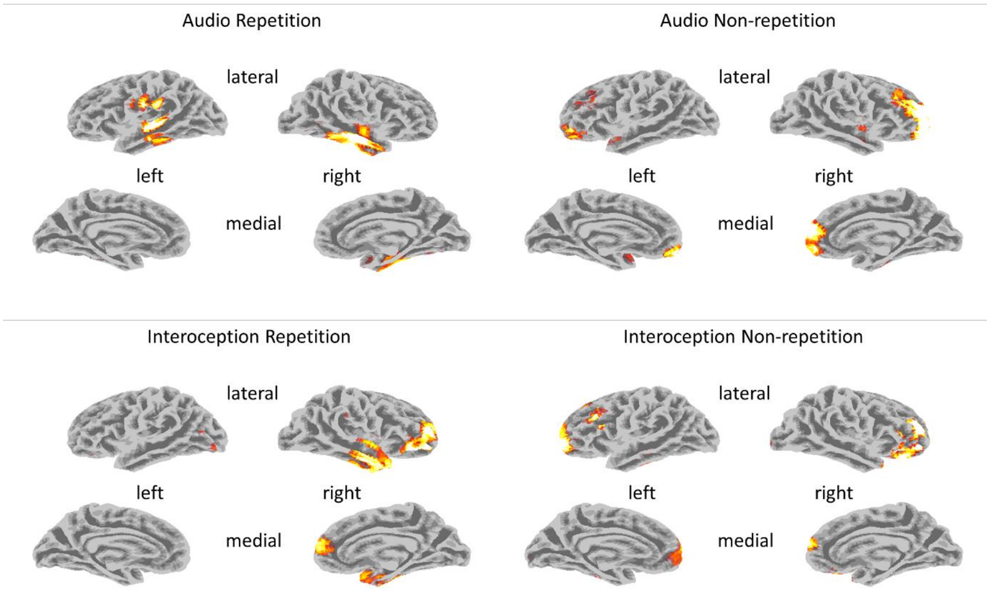


Figure 5.30. [Top] Examples of results of source estimation within gamma band (30-70 Hz) in Korean by eLORETA for the audio repetition (Left) and audio non-repetition (Right). [Bottom] Source estimation for the interoception repetition (Left) and interoception non-repetition (Right).

5.7.4.2. English

In gamma band analysis, the neural genrators for the audio condition (Figure 5.31) were calculated as the parts of right temporal cortex (BA 21, 41, 42; auditory / autitory association cortex), right prefrontal cortex (BA 10) and left temporal pole (BA 38). The temporal pole (BA 38) is related to several cognitive functions such as visual, semantic and socio-emotional processing, and particularly the left temporal pole is known to be interconntected to Broca and Wernicke's area and participate in several language processings such as naming (Ardila, Bernal & Rosselli, 2014). While the sources in the right hemisphere generally overlaps with that of Korean, the activation of the left temporal pole was absent in Korean data. For interoception, the parts of left and right temporal pole (BA 38), right inferior frontal gyrus (BA 45) and left prefrontal cortex (BA 10) were estimated as sources (Figure 5.31). Compared to Korean, where the sources were mainly located in the right hemisphere, the activation in the left temporal pole was dominant in English. Overall, the sources for English were less consistently estimated across the four methods even though it was possible to specify the commonly estimated sources, and estimated sources were more dispersed over several areas of smaller patches. In the low frequency band, similar to Korean, parts of the medial temporal lobe (BA 20, 28, 34, 37) were estimated as sources in both modalities (Figure 5.29). On the other hand, in the analysis for modality within each repetition conditions, four methods did not consistely specify a specific source.

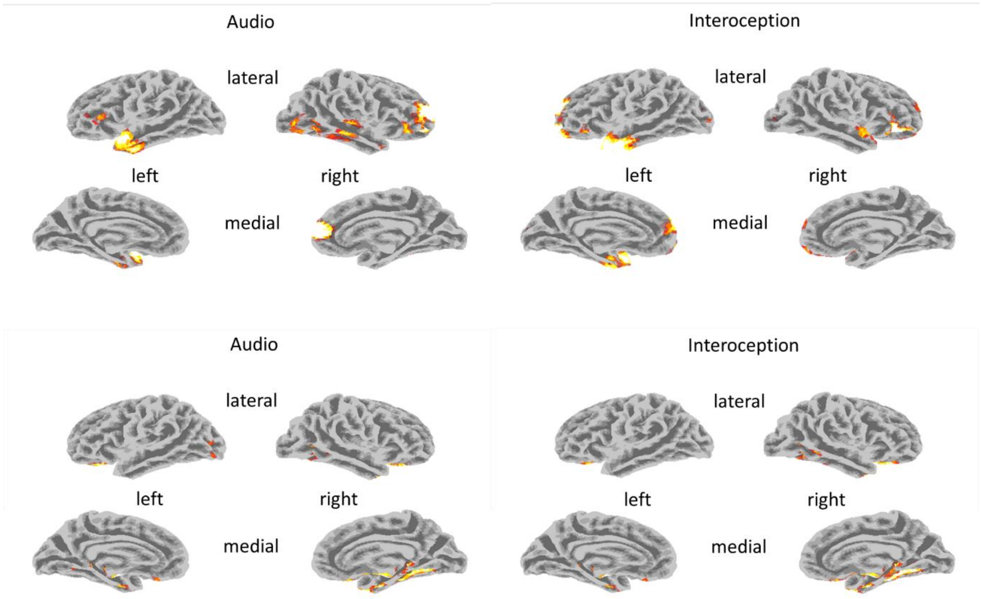


Figure 5.31. [Top] Examples of results of source estimation within gamma band (30-70 Hz) in English by eLORETA for the audio (Left) and interoception (Right). [Bottom] Source estimation within low frequency band (0.1-30Hz) for the audio (Left) and interoception (Right).

To summarize, source estimation in English partly overlaps with source estimation in Korean, but several differences were also found. As in Korean, the neural generator for the audio condition was estimated as parts of the right temporal cortex, and interoception was as parts of the right temporal pole. However, for both audio and interoception, the left temporal pole was calculated as a source, which was not identified in Korean.

5.8. Discussion of Experiment 2

Experiment 2 was to investigate whether the sensory-modality specific word categorization revealed in Experiment 1 was supported by evidence of the actual neural responses and whether there was a difference between Korean and L2 English. The findings of Experiment 2 revealed several differences of neural responses between audio and interoception modality in both Korean and L2 English, partly supporting the perspective of the embodiment cognition. Both in Korean and English, ERP was more positive from the early time window at midline sites and the time-frequency analysis revealed the difference of gamma band power between audio and interoception. Moreover, the estimated sources were different between the audio and interoception conditions, each of which was the region associated with each modality. However, these modality-specific responses were not entirely consistent between L1 Korean and L2 English. Specifically, there was repetition enhancement only in Korean, and the gamma power of audio increased compared to interoception in Korean, whereas it decreased in English. Moreover, in English data, several sources were additionally identified which were not estimated in Korean. In the section that follows, the findings will be discussed further in detail, focusing on repetition effects, modality specific-responses, embodiment of restricted lexical processing, language differences, and regarding issues of embodiment cognition theory.

5.8.1. Repetition effects

The current study found evidence of the repetition effect in both Korean and English, as demonstrated by both behavioral and ERP responses. In the lexical decision task, although there was no difference in accuracy, participants in both Korean and English showed faster response times in the repeated condition. The absence of a repetition effect on can be attributed to the relatively limited impact of masked repetition priming on conscious, offline tasks (Carr & Dagenbach, 1990) and the high level of accuracy exhibited by participants, with an average of approximately 99%, which may have hindered the detection of any potential repetition effect. In addition, ERP responses in both Korean and English revealed a repetition effect in the time window of 100-250 msec, characterized by a more positive voltage (larger P2) in the repetition condition, as well as larger negativity (larger N400) in the repetition condition in the time window of 250-400 and 400-550 msec over central and posterior areas.

Although the repetition effect is not the primary objective of this study, it is still of interest in that it suggested the experiment was conducted properly and the findings from it were not arbitrary. The repetition effect has been found in a number of behavioral and neurological experiments, which is known to lead to a decrease in reaction time and a characteristic neural response such as P2 and N400 (Adelman et al., 2014; Holcomb & Grainger, 2007; Kiefer, 2005; Misra & Holcomb, 2003; Schweinberger et al., 1995). This study also observed that the repetition conditions were associated with shorter reaction times in lexical judgment tasks as well as distinct neural responses—larger P2 and smaller N400 for the repetition conditions. In particular, these neural responses were similar in terms of timing and topography, with that in Trumpp et al. (2014), where the sounds were used which have in common with the material of the present experiments. In summary, the repetition effect of this study is consistent with that of earlier investigations, indicating that the experiment was carried out as designed and that the purpose of introducing the masked-repetition prime paradigm, to measure the effect of unconscious processing, was accomplished.

5.8.2. Modality-specific neural responses between audio and interoception words

The main research objective of this study is to explore whether modality-specific responses exist between audio and interoception and, if so, whether they are also found under restricted lexical processing. First, both Korean and English results showed modality-specific responses between audio and interoception. Although there was no difference in accuracy or response time between modalities in the lexical decision task, ERP analysis revealed different neural responses between them. In both languages, audio elicited more positive ERP than interoception at midlines from early time windows. Although this difference did not seem to be robust because it was not confirmed by non-parametric cluster-based permutation tests, there is no reason to exclude the findings from ANOVA analysis as arbitrary results because it persisted for a considerable amount of time, reaching statistical significance in all three time-windows and resembling the forms of widely known ERP components such as P2, N400 and LPC. Therefore, the difference in neural responses between two modalities can be primarily viewed as evidence in the support of the embodied cognition perspective of language processing.

In the early time window of 100-250 msec, both Korean and English showed significant differences in modalities at midlines sites. The waveforms have a positive deflection which peaks around at 150 msec, following the first negative peak about at 80 msec, forming a typical pattern of the N1-P2 complex. P2 represents early perceptual processing in general, but it is also modulated by other cognitive factors. In language processing, it is known that the amplitude and latency of P2 are affected by various factors such as orthography, length, repetition, frequency, close probability, emotionality, and word category (Hsu et al., 2009; Kanske & Kotz, 2007; Lee et al., 2012; Penolazzi, Hauk & Pulvermüller, 2007; Preissl et al., 1995). Among these, orthography and length are perceptual properties such as shape or size of word form that modulate P2, while frequency, close probability, emotionality, and word category are non-perceptual factors. Two possible reasons can be estimated why non-perceptual factors have influence on the EEG response of the early time window when perceptual recognition of word form is mainly concerned. One is that the prediction of a word (form) facilitates or inhibits the perceptual processing itself (“The word recognition hypothesis”) (Nieuwland, 2019). For example, in the contexts where a specific word is highly expected (high cloze probability), the form of that word will be predicted as well, which will accelerate perceptual processing. The other is that some aspect of semantic processing is performed so rapidly (as soon as it is perceptually recognized) that it intertwines with the perceptual processing. For instance, words with strong emotionality (e.g. taboo words) might affect the neural responses in the early time window as their semantics instantly being processed.

In this study, audio elicited larger P2 compared to interoception both in Korean and English. Since audio and interoception words were matched for word length, frequency, familiarity and word category, it was not likely that P2 was modulated by differences in perceptual properties or differences in predictability (frequency or familiarity) between the two groups. Rather, it might vary due to other factors than those that affect the rate of retrieval of semantic information. In other words, differences in modalities led to different rates of semantic retrieval, which altered the amplitude of P2. In this case, semantic retrieval of the audio for any reason may be sufficiently faster to affect perceptual processing, resulting in a larger P2. This finding may indicate that modality-specific neural responses emerged in language processing early enough to influence perceptual processing. If it were case, it can also be evidenced against the claim that modality-specific neural responses are merely by-products of language processing.

However, this interpretation should be approached with caution, as it requires reasonable account or empirical evidence for why semantic information of audio words was processed sufficiently fast to have influence on early perceptual processing. Only a few semantic elements have been known that influence visual P2. Most researched is the effect of the emotional word on early word processing. Studies on the early processing of emotions have been widely conducted using non-verbal stimuli such as faces and pictures, and it is reported that the processing of emotions can occurs even as early as 100 msec (Kissler et al., 2006). In this respect, research on the early processing of emotional words has an empirical basis as an analogy to emotional object. Also, considering words that cause strong emotional arousal, such as taboo words, it can be reasoned that emotionally charged words is easier to be processed compared to ordinary words, because it focuses more on the instant transfer of emotional state than on the comprehension of meaning. Nevertheless, the effect of the emotionality on P2 is rather limited. The effect is task-dependent, occurring only in more attended tasks, it is found only in the words with positive valence but not in the negative words (Kanske & Kotz, 2007). In other words, although there are reasonable justification and empirical evidence to infer that emotionality affects P2, its resulting neural responses are quite limited and different from research to research. Considering that even the modulation of early perceptual processing by the influence of emotionality is inconsistent, it seems rather unreasonable to merely conclude that audio is a factor that induces sufficiently faster semantic processing without any experimental basis or rational justification. In order to justify this finding, support from neurological research would be necessary, such that auditory words are processed more rapidly or lead to greater P2.

One of the alternative explanations for the greater P2 of audio is that it was a result of the greater N400 of the interoception condition. Nieuwland (2019) suggested that the P2 effect by predictability (Lee et al., 2012) might be due to the influence of early N400. In this study, similarly to Lee et al. (2012), larger P2 of audio in both English and Korean was followed a smaller N400 of audio (larger N400 of interoception). Since N400 starts from 200 msec at the earliest, there may be a temporal overlap with P2 occurring between 150 and 250 msec. Moreover, since the general topography of P2 (the centro-frontal and the parieto-occipital) and N400 (the parieto-occipital) overlaps to a large extent, it is not easy to distinguish the temporal boundary between the two components when polarity difference of two conditions sustains. Therefore, what appeared to be the effect of the P2 may be mixed with the early onset of the N400. This provides a more straightforward interpretation. N400 is a component related to the processing of semantic memory and is assumed to be processed in a distributed manner over various brain regions (Kutas & Federmeier, 2000). Accordingly, from the perspective of embodiment cognition, interoception and audio were activated in different areas, so different degree of N400 was evoked between two modalities. Since N400 amplitude is sensitive to the ease of retrieval from semantic memory or long-term memory (Kutas & Federmeier, 2000), interoception had less ease of access compared to audio, which was reflected in the larger N400.

The assertion that interoception is less accessible than sound in terms of semantic memory retrieval requires further neurophysiological evidence but at least it can be inferred through a reasonable justification. While the source for auditory sense is typically estimated as the auditory cortex, the source of interoception is more distributed from the insular and anterior cingulate cortex to somatosensory cortices (Craig, 2003; Wilson-Mendenhall et al., 2019). Recollecting semantic information from several different regions may be more complicated than from a few concentrated areas. In a similar vein, the low exclusivity of interoception also supports this argument. Exclusivity is a measure of the degree to which a word is experienced as a single sense, and the lower the score, the more multidimensional the word is. Despite attempts to balance exclusivity scores, the exclusivity of interoception words was lower than that of audio due to the intrinsic characteristic of interoception, which is generally more multi-dimensional compared to audio. Multidimensional interoception words should be retrieved through activation of several more regions, which may lower ease of access.

One thing to note here is that ease of access to semantic memory is not synonymous with cognitive processing difficulty. Since N400 is traditionally known as an index of sematic violation, it is assumed to be simply correlated with the difficulty of semantic processing; yet, in practice, N400 is frequently dissociated from response times (Kutas & Federmeier, 2011). For example, concreteness words tend to elicit a larger N400 than abstract words even though they show faster response times on behavioral measures. In this study, no difference in response times between the two modalities was found in the lexical decision task as well. Therefore, the difference in N400 between modalities should not be simply equated with cognitive difficulty but has more to do with semantic memory accessibility.

In the time window of 250-400 msec, where semantics of language is known to be processed, larger N400 for interoception was found in both languages. As mentioned above, this effect continued from the previous time window. As N400 is related to semantic information retrieval, this difference in N400 may suggest the activation of distinct regions for the two modalities in terms of embodied cognition. Since N400 is a complex component influenced by a variety of factors such as frequency, familiarity and predictability, several alternative interpretations can be provided. Although for most of these factors the two conditions were balanced, several factors differed between the two groups, which had influence on N400 according to the previous studies—concreteness and emotionality. As for concreteness, in general concrete words evoke a larger N400 than abstract words (Barber et al., 2013; Kanske & Kotz, 2007; Zhang et al., 2006). Because the perceptual strength of interoception is highly correlated with abstractness, concreteness ratings were not balanced across audio and interoception words, so that audio had a significantly higher concreteness rating than interoception. In this study, however, audio words evoked smaller negativity. Moreover, the additional analysis on concreteness effect[12] did not find a difference in amplitude at midline sites between concrete and abstract words. Regarding emotionality, high-emotional words generally elicit smaller N400 than neutral words (Kanske & Kotz, 2007; Wang, Shangguan, & Lu, 2019). Although the emotional valence between audio and interoception words was not measured, interoception words are likely to have greater emotional valence, because interoception plays an important role in emotion processing (Connell et al., 2018; Critchley & Harrison, 2013) and has a high correlation with emotion (Villani et al., 2019). Furthermore, interoception words in this study included several emotional words such as *afraid*, *angry*, *bored*, *depression*, and *glad*. Nonetheless, actual data demonstrated the opposite trend, with interoception inducing larger N400. Therefore, the N400 effect of this time window appeared to be the result of differences in modality rather than concreteness or affectiveness.

In the time window of 400-550 msec, larger positivity was found in the audio conditions for both languages. This positive-going wave has similar timing and topography with LPC. LPC is assumed to be associated with explicit memory recognition (Rugg et al., 1998). Since this study performed a lexical decision task, participants were required to determine if a presented item was a word or a non-word. Considering the participants' average reaction times (674 msec for Korean, 863 msec for English), this time-window was shortly before judgment for lexical decision was made. In this window, they may attempt to recollect the memory for

[12] The words were divided into two groups based on their concreteness ratings, regardless of modality, and the difference between the two groups was additionally analyzed. This was excluded from the report, however, because the material was not chosen based on concreteness in the first place.

the words more explicitly in order to make lexical judgments. The amplitude of LPC is known to be influenced by the recall of episodic details of memory or the perceived strength of memory (Bechtold et al, 2019; Yang et al., 2019). In other words, the more detailed the memory and the more strongly it is perceived, the greater the positivity. This larger LPC for audio can be justified through the similar assumption as the larger N400 for interoception. The audio conditions, in which memory was recalled in a less distributed manner, were likely to be perceived more strongly than interoception. However, since LPC is often task-dependent and related with various cognitive factors such as task difficulty (Kopf et al., 2013), it may have been affected by a variety of factors not examined in this study. Moreover, since the lexical decision task is not the one that directly or explicitly requires memory recall, the memory recognition process might not be comparable to explicit memory recognition generally dealt with in LPC research, even though the modulation of LPC by enhanced retrieval was reported in a lexical decision task as well (Bechtold et al, 2019).

A modality-specific response was found in both languages not only in the ERP analysis but also in the time-frequency analysis. In Korean, audio had greater gamma power than interoception in the early time window over posterior regions. In English, there was a difference in the gamma band as well, but in contrast to Korean, the power for interoception increased in the later time window over central regions. In other words, there was a difference in the modality of the gamma band in both Korean and English, but the difference was in the opposite direction. Gamma-band synchronization is found across numerous cortical regions, induced by various stimuli or tasks, and associated with a variety of cognitive abilities (Fries, 2009; Hipp et al, 2011). One of functions related with it is conscious perception. Gamma-band power is known to increase in response to consciously perceived auditory stimuli and related with the activity of auditory cortex (Leicht et al, 2021; Morillon et al., 2010; Pantev, 1995; Steinmann et al., 2014). The relatively increased gamma band power of audio conditions in Korean may be explained as the activity of the auditory cortex from the perspective of embodiment cognition. However, since gamma band frequency is related to perception of other senses as well as audio, it cannot be simply interpreted as an increase in audio perception (Karns & Knight, 2009; Mori et al., 2013; Zhang et al., 2012). A speculative conclusion that can be derived from the current data is that there was a difference in the neuronal connections between audio and interoception words because the gamma band frequency reflects the neuronal oscillation between diverse regions. This may imply that audio and interoception words activated different parts of the brain.

On the other hand, in English, the gamma power in the audio condition rather decreased. Nonetheless, this should not be simply interpreted as the non-involvement of auditory cortex, because the reduction of gamma power persisted slight longer to a later time window in English. Since the neural oscillation associated with sensory perception typically occurs shortly after a stimulus onset (Pantev, 1995), the gamma band decrease in the later time window of English may not be related to auditory perception. In the early time windows of English, the gamma band power of audio condition was greater than the baseline, which was comparable to Korean. The opposite trend of the gamma power difference in modalities between Korean and English seemed to be more attributable to differences in interoception. In Korean, the gamma power of interoception was generally decreased and smaller than that of audio, whereas in English, it was enhanced over several regions and appeared to be greater than that of audio. One possible account for this is that the gamma oscillation had to do with the retrieval of memory. A study on the relationship between working memory and gamma band revealed that gamma power increased when retrieval-related burden was present (Tallon-Baudry et al., 1998). In this respect, the gamma increase in the interoception condition may indicate that its memory retrieval was highly active or challenging compared to audio. In the lexical decision task, the response times of interoception although not statistically significant, were longer than those of audio in English. This may suggest that L2 requires more memory retrieval for interoception processing compared to L1. However, this difference in interoception was not statistically significant; so, interpretation must be undertaken with caution.

What is crucial from the standpoint of embodiment cognition is whether the responses of the two conditions actually had different neural generators, as in the above interpretation that the difference in ERP and time-frequency analysis was attributable to the activation of distinct regions. The results of source estimation partly supported this interpretation. In Korean, audio and interoception were estimated to have disticnt neural generators. Parts of the auditory cortex, auditory association cortex, and dorsolateral prefrontal cortex were considered to constitute audio sources. Among them, the right auditory and auditory association cortex were regions particularly distinguished from the source of interoception. They are associated with audio perception, and thus expected to be activated when processing audio words according to embodiment cognition theory. The sources of interoception was estimated from the right temporal pole, right medial orbito prefrontal cortex, right inferior frontal gyrus, and left prefrontal cortex. In particular, the right medial orbito prefrontal cortex and right inferior frontal gyrus differed from the audio condition. The right inferior frontal gyrus is connected to the insular and is known as related to interoception. The left inferior frontal gyrus is part of Broca's region, which is assumed to be associated with speech production, was not estimated as a source. The rest of the estimated regions are all known to contribute to complex cognitive functions. One of their functions is socio-emotional processing, which is claimed to be related to the embodiment of abstract concepts, according to Words as Social tools theory (Borghi et al., 2019). The finding from the source analysis of this study seemed to support this claim as well. However, since ERP analysis did not find the effect of emotionality, further examination is necessary to in order to interpret the activation of these areas in relation to socio-emotional processing.

In English, distinguished brain sources were also observed between the two modalities. The sources of audio were parts of right auditory and autitory association cortex, right prefrontal cortex and left temporal pole; the sources of interoception were calculated as parts of left and right temporal pole, right inferior frontal gyrus and left prefrontal cortex. Estimates of each condition include areas related to audio (auditory cortex) and interoception respectively (inferior frontal gyrus). The estimated area was generally consistent with Korean, but the degree of agreement between approaches was generally lower than that of Korean, and the estimated sources were more dispersed in small patches. This may be owing to the low signal-to-noise ratio, but since the experiment was designed as a within-subject comparison, there would be no reason for it to be lower than that of Korean. This may rather reflect the characteristics of the L2 processing. One possible explantion for this is that the sensory experiences of each participant varied, causing the activated regions to be spread and making focused estimation more difficult. In fact, it is known that the networks involved in L2 processing differs more greatly among participants compared to L1 processing (Dehaene, et al., 1997). The other is that since overall quality and quantity of the sensory experience in L2 were relatively poorer, each source was less activated and the strength of the source estimates was weakened, resulting in more distributed estimation. Finally, it may be because L2 processing necessitated participation in other regions that were not strongly activated in L1 processing. It is difficult to determine which of these interpretations was most accurate based on the available data. Perhaps this finding was a combination of all these interpretations. However, these implied that the participants in L2 had a different embodiment from L1—embodiment was weaker, or varied across participants, or invovled more regions.

The sources of English almost overlapped with that of Korean, but several regions were additionally estimated, one of which was the left temporal pole. Although the role of the temporal pole is not well explored, it is regarded as semantic hub in several theories (Patterson et al., 2007; Visser et al., 2010) and also known to be involved in many complex higher-order cognitive functions such as visual processing, semantic processing in all modalities, and socio-emotional processing (Herlin et al., 2021). In particular, the left temporal pole seem closely related to a variety of language processing functions, including semantic processing, speech comprehension, naming or word retrieval for specific entities, and the processing of abstract information, and interconnected to Broca and Wernicke's area (Ardila et al., 2014). There is no evidence that the left temporal pole is specifically unique to L2 participants' neural responses, but rather, low proficient learners showed lower activation in this area than high proficient learners or L1 speakers (Perani et al., 1998). Since in this study the sources were estimated on the basis of relative strength, no or less activation of the left temporal pole in Korean data did not suggest that it was not involved in L1 processing, but that it is relatively less active compared to other modality-specific domains. In other words, in L2, the modality-specific region and the left temporal pole were activated at relatively similar levels, in contrast to L1.

The findings that the left temporal pole was additionally activated only or relatively strongly in L2 processing can be interpreted in several ways. The temporal pole and the anterior temporal lobe around it are considered as one of the areas related to amodal semantic hubs, where semantic processing occurs regardless of the modality of words. Therefore, activation of this region may imply disembodiment of L2 processing. However, as described above, L2 processing demonstrated nearly similar modality-specific neural responses to L1 processing. Thus, it seems difficult to conclude that L2 processing is disembodied differently from L1 only with the activation of the left temporal pole.

An alternative interpretation for the activation of the left temporal pole is that L2 processing was less embodied compared to L1. Some hybrid-approaches of embodiment theory argue that in contrast to concrete concepts, abstract concepts do not have a modality specific ground, so they are processed in an amodal semantic hub. This interpretation was partially refuted with the findings that the left temporal pole was equally activated for both modalities even though the audio words were relatively more concrete. Under these assumptions, processing of more concrete, audio words should have less or no activation of these regions. However, it can be argued that some concrete concepts for L2 were processed as in abstract concepts for L1 because L2 processing had no or weak embodied association for some concrete concepts. Audio words were relatively more concrete than interoception words, but their concreteness ratings were only slightly above average. In L2 processing, perhaps only extremely concrete words are processed in modality specific regions, while others are processed in the amodal hub. According to this interpretation, the activation of the left temporal pole in both modalities may imply that both audio and interoception words were processed as abstract concepts in L2 processing, i.e., less embodied. However, this interpretation requires further explanation of why the left temporal pole was not estimated as the source in L1 interoception condition. Although, as stated above, no estimation does not mean the absence of the activation, stronger involvement of the amodal hub should have been observed in interoception words, which were relatively more abstract. In order to explain the data of L1, additional assumptions such that only extremely abstract concepts are processed in amodal hub in L1 processing or L1 is qualitatively different from L2 so that the amodal hub is not used will be needed.

Another alternative interpretation is that the embodiment of L2 processing was more dependent on linguistic experience. WAT emphasized the role of language and sociality in the embodiment of abstract concepts (Borghi et al., 2019). According to the WAT, as there is no specific referent, abstract concepts are acquired primarily through language or by relying on others to figure out the meaning; hence, they are embodied through linguistic and social experience. This assertion pertains mostly to the acquisition of abstract concepts for L1 speakers, but it may also be extended to the acquisition of concrete concepts for L2 speakers. It is likely that both language and sociality are crucial for the acquisition of concrete and abstract concepts in an L2 learning environment. In particular, the effect of language experience is more likely to be vital, because in many cases, L2 word are explicitly learned through L1 translation equivalents. In fact, in L2 lexical processing, parallel activation of L1 translation equivalents is frequently reported (Thierry & Wu, 2007); many bilingual mental lexicon models also presume that the lexicons of L1 and L2 are interconnected (Dóczi, 2019). In this regard, the activation of the left temporal pole may indicate that embodiment via language experience plays a more crucial role in L2 processing. Moreover, this interpretation is compatible with the absence of the activation of the left temporal pole in L1 processing. WAT does not assume that all abstract words are related to linguistic experience, but those with a lower level of emotionality and perceptual strength. Therefore, linguistic experience is not necessarily utilized in processing audio or interoception words in L1 processing. In this respect, among several possibilities, the interpretation that the estimation of the left temporal pole in L2 processing was derived from its greater reliance on linguistic experience for embodiment seems most appropriate.

While the sources at higher frequencies were associated with modality-specific regions, at lower frequencies, the medial temporal lobe, a region related to the hippocampus, was calculated as a source for audio and interoception both in Korean and English. The medial temporal lobe is known to play an important role in memory input and retrieval (Meyer et al., 2005) and regarded as the biological base of declarative memory in the procedural-declarative memory system (Raslau et al., 2015). As one of multiple sources of N400 (Kutas & Federmeier, 2011), it is also acknowledged for its role in semantic integration (Meyer et al., 2005). These findings might suggest that modality-specific neural responses mainly occur in higher frequency bands, while responses for general cognitive processing appear in lower frequencies. Given that the gamma band is related to conscious sensory perception and the study found a difference between modalities in this range in the time-frequency analysis, different estimated sources for each frequency band is compatible with the theoretical background and other findings of the current study.

While the aforementioned interpretation is intriguing, it must be approached with caution. Previous research has indeed investigated the functions associated with different frequency bands in language processing. For example, Lam et al. (2016) found that specific frequency bands appeared to be more closely associated with certain language processing tasks (e.g. the theta band with lexical retrieval) and involved the activation of different brain regions. Therefore, the finding that different sources were estimated according to each frequency is not unexpected. However, it must be noted that a consensus on the function of each frequency band has yet to be established. Some studies suggested that various functions were associated with similar frequency bands, while others indicated that the same functions were related to different frequency bands (Prystauka & Lewis, 2019). Moreover, there are limited studies that demonstrate the correlation between the functions of each frequency band and the activation of specific neural generators (Lam et al., 2016; Moreno et al., 2015). Furthermore, EEG source localization has the relatively lower resolution for medial parts of cortices, such as the medial temporal lobe.

5.8.3. Modality-specific responses in limited lexical processing

The next question of this experiment was whether these modality-specific responses also appeared in limited lexical processing. One of the key issues in embodiment cognition theory is the discussion of whether embodiment processing is essential to language understanding or is a mere by-product. This study dealt with this issue by examining whether modality-specific responses would emerge even in restricted lexical processing through the repetition effect. Typically, the repetition effect reduces the reaction time and decreases or increases the neural responses depending on the type of repetition. When the number of the repetition is fewer and it has low visibility, the repetitions generally enhance neural activities (Müller et al., 2013; Turk-Browne et al., 2007). Thus, the current experiment's paradigm is more likely to generate the repetition enhancement than repetition suppression.

First, no interaction between modality and repetition was identified in the analysis of the response time in the lexical decision task, which indicated that the repetition effect was not different across modalities. Although particularly in English, the decrease of the response times by the repetition was numerically greater in the audio condition, but the effect was not significant. This might suggest that there was no modality-specific response in restricted lexical processing. Differences in response time reduction by the repetition might be caused by different characteristics of audio and interoception in the early automatic processing. For instance, audio may be activated faster and more strongly in a brief period of time, or it may only require shallow activation for lexical processing. Or it may be because interoception words have relatively incomplete lexical processing in a limited range of activation. However, words that are easily processed may not have this reduction effect in the first place, as high-frequency words show rather smaller repetition effect than low frequency words (Kinoshita, 2006). Thus, even if there was a difference in modality-specific response in limited lexical processing, it may not result in significant variation in response time, because that difference was not large enough to affect behavioral responses.

On the other hand, in the ERP analysis, differences in the modality-specific responses by the repetition effect were partially confirmed. First, in Korean, there was an interaction between Modality, Repetition, and Hemisphere in the time-window of 100-250 msec. This may be because the repetition effect occurred only within the audio conditions particularly in the left hemisphere, although it was not statistically confirmed. More importantly, when the data were divided according to each repetition condition, the analysis of separate datasets revealed a difference in modality-specific responses between the repetition and non-repetition condition. In the repetition condition, a significant difference between modalities was found at midline sites, but not in the non-repetition condition. Under the assumption of "repetition enhancement," this can be interpreted as each modality-specific response being stronger as a result of the repetition, and the distinctions across modalities that were not fully revealed in the non-repetition condition becoming more pronounced in the repetition condition. In other words, the region activated by the masked audio words was activated stronger due to the repetition, while the interoception region was also further activated by the repeated interoception words, which highlighted the difference in neural responses between the two modalities.

This interpretation seemed to be consistent with the results of source estimation for the modality within each repetition condition. The estimated source in the audio repetition condition was concentrated in the left and right temporal lobes which includes the auditory cortex, whereas the source in non-repetition was mainly estimated as part of the pre-frontal cortex. But it does not mean that activation of temporal lobe was absent in source estimation for the non-repetition condition. In the non-repetition condition, the temporal lobe was also assumed to be part of the source, although the estimated strength for it was different across methods and in general the strength was relatively weak. Particularly, when the source estimation threshold was lowered (95% → 90%), the activation around the temporal lobes was more prominent. In other words, in the non-repetition conditions, several regions including the temporal lobe were activated, but it seems that the activation was stronger in the pre-frontal cortex than in the temporal lobe. In contrast, the temporal lobe activity was most apparent in the repetition condition. In this regard, the following assumptions can be made. For the processing of masked primes of the audio condition, audio-related regions such as temporal lobes were activated, and the activation in these regions was enhanced by repeated targets. On the other hand, in the non-repetition condition, because the temporal lobe was less activated while other regions were more activated in the processing of the unrelated masked prime, there was no enhancement of the temporal lobe in the non-repeated target, and its activation was rather modest.

Similar repetition enhancement was observed in the interoception condition, but not as distinguished as in the audio. Partial activation of the inferior frontal gyrus, which is regarded to be related to interoception, was observed in both repetition and non-repetition conditions. In the repetition condition, the right temporal pole was more strongly activated compared to the non-repetition condition. Since it is involved in complex cognitive functions, its activation does not always suggest a modality specific response of interoception. However, since this area was consistently engaged in the non-repetition condition and the English data, it seemed to be associated with interoception words. In this respect, relatively strong activation of the right temporal pole may imply repetition enhancement. In contrast, the region that was relatively more involved in the non-repetition condition was the prefrontal cortex, as it was in the audio non-repetition. Given that the pre-frontal cortex is related to execution functions such as attention, the activation of this area may reflect the attentional switching cost for processing the target that do not match with the prime. In summary, the repetition enhancement occurred in Korean, and the resulting enhanced regions for each modality were mostly modality-specific regions. This may demonstrate that modality-specific responses were present even during the limited lexical processing.

In English, similar to Korean, there was an interaction between repetition and modality in the early time window. In the time window of 100-250 msec, a repetition effect was greater in interoception over the anterior regions but in audio over the posterior regions. It seemed that the topography of the repetition effect was different between the modalities in the early time-window. However, since this trend was not clearly supported by post-hoc analysis or the previous studies, it is difficult to discuss the difference in modality-specific response through this finding of the interaction. Additionally, in contrast to Korean, when the repetition and non-repetition conditions were analyzed separately, similar modality-specific neural responses were revealed in both conditions. They were not completely identical, but no noticeable difference between them was found. One difference was that the differences between the modalities appeared to be distributed wider in the non-repetition condition, as significant differences were found in the right hemisphere as well as in the midline sites. Furthermore, it was not possible to estimate a specific source in each of the repetition condition through source estimate analysis. The estimated sources were less consistent across the analysis methods and they were more dispersed in forms of smaller patches. A variety of factors such as the heterogeneity of L2 learners contribute to this failure, as described above. Moreover, the data being divided into the repetition and non-repetition conditions, the number of the samples were reduced along with the signal-to-noise ratio, which possibly made estimation more difficult. In short, in English, no clear evidence was found for changes in modality-specific responses by the repetition effect.

Two possible explanations can be put forward to account for the lack of repetition enhancement observed in English. The first interpretation is that the duration of the masked prime was too short for the L2 processing, preventing sufficient lexical processing from occurring. In order to maintain identical conditions for Korean and English, the duration of masked primes was kept the same for both groups, which may have been too brief for L2 processing. Although as demonstrated by the successful repetition effect, the masked prime was recognized even in the L2 condition and had an unconscious influence on the processing of the target word, it should be noted that the repetition effect is not necessarily caused by semantic processing. Since this study used the repetition prime task that presented the same word in the prime and target, the repetition effect could occur for orthographic and phonological reasons without a semantic access to words. To further investigate this interpretation, future studies may need to increase the length of the masked prime.

Another interpretation is that limited lexical processing fails to activate some modality specific regions when the connection between language and environment is relatively weak. From the standpoint of embodied cognition theory, semantic processing does not occur within a single area, but rather via the coordinated activation of many regions associated to the sensory-motor experience of the word. If the degree of embodiment is poor and the connectivity among several regions is weak, only a few stronger connections are engaged in a constrained cognitive processing. In the case of L2, because the association between environment and language is relatively weak compared to L1, modality-specific regions may not be activated or less activated in the limited lexical processing by masked primes.

5.8.4. Differences between L1 and L2

The results of the experiments indicate that overall, there is a similar degree of embodiment pattern in L1 and L2 processing. The patterns of ERP's modality-specific responses were similar from the early time window in English and Korean. These findings might suggest that L2 processing is embodied in a similar manner as L1 processing.

However, the neural responses between L1 and L2 were not entirely identical. Discrepancies were noted in the repetition enhancement, gamma band power patterns, and estimated neural generators. These differences appear to imply weaker embodiment of L2 processing or the influence of L1's intervention. In contrast to L1, repetition enhancement was not observed in L2. This shows that activation of modality-specific regions does not occur in restricted lexical processing such as masked primes in L2, perhaps due to the weaker embodiment of L2 processing. However, as mentioned above, the absence of repetition enhancement may merely mean that the duration of the masked prime was too short to be recognized in L2.

In time-frequency analysis, there was a difference in gamma band power between modalities in both Korean and English, but Korean and English showed different patterns in their polarity, topography, and timing. In particular, gamma power was increased in the audio condition in Korean but increased in interoception in English. This difference between languages appeared to be due to the increased gamma power of the English interoception condition since the audio conditions of the two languages were similar. Since the gamma band is associated with neuronal interaction, this discrepancy at least showed that there was a different pattern of neural connection required for word retrieval between Korean and English. Since the interoception condition generated the most notable difference between Korean and English, it can be argued that such a difference was attributable to the retrieval of interoception words. Interoception is a less noticeable sense and information from it often does not reach the conscious level (Holzer, 2017); moreover, abstract concepts related to interoception have more heterogeneous referents or experiences across individuals. Because interoception words were not grounded on a specific referent with an easily perceivable sense, the formation of their embodiment might require more interaction with the environment than that of audio words. In this regard, interoception may have exhibited more distinguished neural responses for L2 learners, who have less experiences of associating with sufficient environment.

In source estimation, the audio and interoception conditions, common to both Korean and English, activated the auditory cortex and interoception-related areas. However, there were major differences in two aspects. The first difference is that in English, the consistency of source estimation was poorer, and the sources were estimated in the form of small patches or in a less intensive fashion. This may imply that the embodiment of L2 processing is based on a wider variety of areas. L2 participants were a group with generally similar experiences in language acquisition, usage, and proficiency, but due to the peculiarities of L2 learners, they inevitably have different L2 experiences. So, their embodiment can be grounded in more diverse sensory modality or can vary in degree across individuals, which may lead to the estimation of distributed sources.

The second difference is that the intervention of the left temporal pole was prominent in English compared to Korean. This region was estimated as a source in both audio and interoception in English data. The role of the temporal pole is unclear, but it is known to be related to linguistic, social and emotional processing. Among them, the left temporal pole is assumed to be particularly associated with language processing and connected to the language-related regions such Broca's and Wernicke's area. This area is also considered as (amodal) semantic hubs in amodal theory and hybrid approaches of embodiment cognition. Of the many possible interpretations, the most congruent with the data is the one that activation of the left temporal pole indicates the embodiment of L2 depends more on linguistic experiences, compared to L1 processing. Borghi et al. (2019) argued in WAT that abstract concepts with no specific referent are embodied through the experience of rehearsing their definition verbally. Given the explicit learning environment of L2 learners which is mediated through L1 translation equivalent and L2 mental lexicons that are lexically linked to L1, conscious or unconscious verbal rehearsing is expected to participate in L2 processing regardless of the availability of referents. Therefore, L2 processing is more likely to be embodied through linguistic experience, whether concrete or abstract words, which was reflected as the activation of the left temporal pole.

5.9. Summary of discussions

The issues of embodiment cognition theory will be discussed by summarizing the discussion so far. First, this experiment, under the assumption of embodiment cognition theory, aimed to examine whether words related to specific sensory modalities are processed in the corresponding sensory areas. The findings from ERP analysis, time-frequency analysis, and source estimation showed that there were differences in neural responses between the two modalities, audio and interoception, and the neural generators were estimated as part of regions known to be associated with each modality. These results are consistent with the predictions of embodiment cognition theory.

Next, another important issue of the embodiment cognition theory is whether the sensory-specific neural responses are essential for language processing or a mere by-product of initial semantic processing. In this study, this was explored through the masked repetition effect. The results revealed that in L1, modality-specific responses were also activated in the restricted lexical processing, which was supported by repetition enhancement in the repetition conditions. Although repetition enhancement was not observed in English, the presence of the P2 and early N400 effects indicated that modality specific responses were elicited from an early time-windows in English as well. Although it would be difficult to conclude whether embodiment processing is essential for language processing only with these findings, it is at least demonstrated that modality-specific responses were activated from an early stage of processing and in the subliminal way in L1.

In addition, the current study attempted to explain the embodiment processing of abstract words through a rather unknown sensory modality, interoception. There has long been debate about whether abstract concepts are also embodied, and if so, what they are grounded on. Interoception is a sense highly correlated with abstractness and often regarded as a mechanism to explain the embodiment of abstract concepts. The results showed that the interoception related words evoked distinguished neural responses and was estimated to have different neural generators from the audio-related words. Furthermore, according to the results of the ERP analysis, despite high correlation with abstract words, the interoception words elicited neural responses with different characteristics from typical ERP modulations of abstract words in the previous studies. Although further research on the relationship between abstractness and interoception is necessary, the findings of this study at least showed that the account for the embodiment of abstract words based on the concrete-abstract dichotomy is incomplete. In this respect, rather than simply categorizing words into abstract or concrete, it will be necessary to reclassify abstract words based on sensory characteristics under the assumption that abstract words are grounded on the same base of the sensory-motor areas with concrete words.

At last, this study sought to explore the difference between L1 and L2 in embodiment processing. The results of the experiment showed that the embodiment processing of L1 and L2 was relative comparable in several aspects. In spite of differences in processing speed, patterns of behavioral response and ERP responses did not reveal a clear distinction between two languages. However, two languages yielded different results in terms of the limited lexical processing. In Korean, modality-specific neural responses were enhanced by repetition, whereas in English, the repetition effect did not significantly change the modality-specific neural responses. However, it does not necessarily mean that L2 modality-specific responses are epiphenomenal spreading from amodal sematic processing, since L2 modality specific-responses appeared sufficiently early time windows from the ERP analysis. Another difference between L1 and L2 is that the left temporal pole was additionally estimated as a source in both modalities in L2 processing. As it is related to language processing, its activation may indicate that the embodiment of L2 processing is more dependent on linguistic experience than L1.

Chapter 6. General Discussion

This dissertation conducted two experiments with the goal of answering three general research questions. Research questions are as follows: (i) Is the embodiment of lexical processing essential for comprehension or mere epiphenomenal activation? (ii) Are abstract concepts which lack a specific referent embodied as well? Does interoception play an important role in the embodiment of abstract concepts? (iii) Is L2 processing embodied although it has fewer direct interactions between language and environment? If so, how does it differ from L1 processing?

To answer these questions, two experiments were conducted. Experiment 1 established the perceptual strengths norms of L1 Korean, L1 English, and L2 English were established and compared their distributions. In constructing perceptual strengths norms, by including interoception as one of the perceptual senses, it was examined whether interoception is related to the embodiment of abstract concepts. In addition, by exploring whether L2 users could associate language with sensory perception and comparing how the results of such associations differed from those of L1 users, the influence of the environment on language was investigated. Experiment 2 examined whether audio and interoception words elicited modality-specific neural responses in L1 and L2 processing through EEG experiments. Interoception, a sense related to abstractness, was investigated as one of the possible sensory modalities in which abstract concepts might be grounded. By utilizing the high temporal resolution of EEG and a masked repetition prime paradigm, this study also aimed to determine whether modality-specific responses were essential for semantic processing or simply a by-product.

The results of the experiment were as follows. (i) Through experiment 2, modality-specific neural responses, as evidence of embodiment, was found in both L1 and L2 processing from early time-windows. Particularly, in L1 processing, modality-specific neural responses were revealed even in restricted lexical processing triggered by masked primes. This demonstrated that the activation of modality-specific regions found in lexical processing was not just an epiphenomenal activation. (ii) In Experiment 1, interoception showed a high correlation with abstract concepts and was generally distinguished from other senses. This may indicate that interoception played an important role in the embodiment of abstract concepts that were seemingly not embodied by other senses. Experiment 2 also proved that interoception had modality-specific responses distinct from audio. Audio and interoception induced different neural responses in ERP and time-frequency analysis and were estimated to activate regions related to each sense in source estimation analysis. (iii) Experiments 1 and 2 revealed that there was no noticeable difference between L1 and L2 in the distribution of perceptual strength norms and modality-specific responses. This might suggest that L2 processing was embodied comparable to L1 processing despite relatively fewer contacts with the environment. However, in Experiment 1, L2 processing was influenced by the native language of L2 users, revealing a slightly different distribution from L1. Moreover, in Experiment 2, L2 processing was not completely identical to L1. The topography and polarity of gamma band frequency were different in time-frequency analysis, and the left temporal pole was additionally estimated as neural generators compared to L1. This might imply attenuated modality-specific responses or responses that rely more on linguistic experience in terms of embodiment.

Below, the above-summarized results will be discussed in detail and the theoretical and methodological implications of this dissertation will be suggested.

6.1. Embodiment as an essential lexical processing or epiphenomenal activation

The first research question of this dissertation was whether sensory motor activation is part of lexical processing essential to language comprehension or mere epiphenomenal activation of post-conceptual processing. The first step was to observe the activation of early processing. This would disprove the claim that activation in sensory-motor areas is only a by-product of semantic processing of amodal regions, by demonstrating the presence of embodiment in early cognitive processing before activation spreads. To this end, this study introduced EEG with great temporal resolution as an experimental technique.

The results showed modality-specific neural responses from early processing in all of ERP, time-frequency analysis and source estimation. In the ERP analysis for both Korean and English, the audio condition induced a larger P2 in the time window of 100-250 msec, and the peak of P2 was approximately 150 msec. In the time-frequency analysis, there was a difference in gamma band power between audio and interoception, which occurred from about 280 msec in Korean and about 300 msec in English. In the source estimation analysis, the sources for each modality were estimated in the 140-160 msec in Korean and in the 160-180 msec in English. Given that N400 components, which are closely associated with semantic processing, are evoked as early as 200 msec, it is evident that semantic processing generally begins around at 200 msec. The time window at which modality-specific responses was first found in Experiment 2 approximated that of initial semantic processing. Nevertheless, it is possibly argued that the modality-specific neural responses were the result of the spreading activation from the amodal areas which were activated even earlier than 200 msec. However, the time-window before 200 msec is primarily associated with early perceptual processing, and it seems less convincing that the amodal region is already activated when the physical properties of stimulus are barely detected.

The second approach of this study is to induce lexical processing that suppresses spreading and then examine the modality-specific responses from those restricted activations. This study utilized a masked repetition priming paradigm and measured modality-specific responses by the repetition effect. In Korean, repetition enhancement was observed, which refers to the increased neural activity by the repetition. This particularly occurs when the visibility of the prime is low and the number of iterations is minimal. These results indicate that modality-specific responses in Korean were present even in the cognitively constrained state triggered by masked primes. In contrast, there was no indication of repetition enhancement in English. However, this does not simply mean that the modality-specific responses reported in English were the epiphenomenal activations from amodal regions. As stated above, the modality-specific responses also appeared during early processing in L2. Rather, the lack of repetition enhancement in English may be due to either the masked prime being too brief for L2 processing, preventing essential lexical processing from occurring, or to limited lexical processing failing to activate modality-specific regions as a result of weaker embodiment in L2 processing.

In summary, the modality-specific neural responses were observed from an early time windows, which refutes the claim that sensory-motor activation in language processing is only a by-product of semantic processing of amodal hub. Moreover, since in Korean modality-specific responses were also elicited by masked primes restricting spreading activation, they seem to be essential and concurrent rather than secondary activation. However, repetition enhancement was not revealed in English, which may be evidence of weak connectivity between words and modality-specific regions in L2.

6.2. The embodiment of abstract concepts

In the Literature review, three main groups of ideas on the embodiment of abstract concepts were introduced. The first argument considers abstract concepts as a homogeneous group opposite to concrete concepts and claims that they cannot be explained from the point of view of embodiment cognition (Dual Code Theory; Paivio, 1986). These views either support the amodal theory of language processing or consider such an approach more appropriate at least for abstract concepts. The second view argues that abstract concepts are not completely separated from concrete concepts and are embodied similarly as concrete concepts, but possess several unique properties distinct from concrete concepts. These are including the theories such as the affective embodiment account (Kousta et al., 2011) or the Word As social Tools (WAT) (Borghi et al., 2019). They seek to explain the embodiment of abstract concepts by adding several features such as emotion, social and linguistic experiences for which abstract concepts are regarded to be grounded. A third viewpoint is that abstract concepts are embodied in the same manner as concrete concepts, because abstract and concrete concepts vary only in degree of concreteness or perceptual strength and are not classified into distinct categories. According to this idea, abstract concepts are equally grounded to the sensory motor cortex as are concrete concepts.

This dissertation focused on the third approach but did not rule out the claims made by the second group. The argument put forth by the third group is deemed the most straightforward, as it proposes that abstract concepts can be understood through the existing framework for concrete concepts. However, since the activation of the brain areas related to five commonly recognized senses alone cannot fully account for the embodiment of abstract concepts, this study incorporated interoception as part of the sensory-motor system. Additionally, according to WAT, there is also evidence that regions of the brain related to language and sociality are activated in the processing of abstract concepts. Since the activation of these areas cannot be described only by sensory-motor system, even with the inclusion of interoception, this study also explores the possibility of the second point of view.

First of all, the results refuted the first point of view that abstract concepts are in opposition to concrete concepts and they cannot be explained in terms of embodiment. Experiment 1 showed that words that could have been classified as abstract concepts under the abstract/concreteness dichotomy were distinguishable by various senses. Although both audio and interoception words were closer to abstract concepts having a positive correlation with abstractness, the language users did not treat them as the same group, evidently differentiating them as belonging to distinct sensory modalities. Furthermore, ERP and time-frequency analysis of Experiment 2 revealed modality-specific neural responses between audio and interoception, which were related with abstract concepts. These differences can be interpreted as the ease of memory retrieval (N400) or the difference in neural connection (gamma band oscillation), suggesting the involvement of distinct brain regions between two modalities. In fact, source estimation analysis estimated the sources of the two modalities differently, and the estimated sources were parts of the regions related to each modality. As a result of demonstrating that abstract concepts can also be explained in terms of embodiment, it supported the assumption of embodiment cognition theory.

Previous studies with the first perspective have classified abstract concepts into a single group opposite to concrete concepts and reported the disembodiment of abstract concepts. However, those findings most likely resulted from mislabeling or misgrouping of several different abstract concepts. Previous studies have often labeled words that can be categorized into several different sensory modalities as abstract concepts and contrasted them with concrete concepts (Dalla Volta et al., 2014; For more details in this issue, see *2.1.3. The embodiment of abstract concepts* in *Literature review*). However, as shown in the results of Experiment 1 and Experiment 2, abstract concepts are unlikely to belong to the same category. Therefore, in order to more accurately assess the embodiment of abstract concepts, a more detailed classification is required which takes heterogeneity of abstract concepts into account rather than a simple abstract/concrete dichotomy. If abstract concepts are classified based on appropriate criteria, then the embodiment of subgroup of abstract concepts will be observed. This research viewed categorization based on the neuroanatomical basis of sensory-motor regions as a suitable criterion. Audio and interoception words classified by this classification elicited different neural responses and activated different regions, in accordance with the predictions of embodiment cognition theory.

Nevertheless, the findings that audio and interoception were distinguished by perceptual strength norms and elicited different neural responses is not a complete refutation of the abstract/concrete dichotomy, because audio was composed of words with higher concreteness than interoception. Since interoception is highly correlated with abstractness, interoception words inevitably have higher abstractness than audio words. Therefore, the difference between audio and interoception could be interpreted as the same as the difference between concrete and abstract words. However, the results of the ERP analysis contradicted the generally observed concreteness effect. Concrete words are known to be processed faster and have a larger N400, but in this study, audio words with high concreteness did not show shorter responses time nor larger N400. Therefore, the difference between audio and interoception cannot be interpreted as a concreteness effect reported in the previous studies. Rather, it is evidence against the concreteness effect, refuting the assertion that abstract concepts are categorically distinguished from concrete concepts.

It was unclear, however, whether the second or the third argument was favored by the findings. Both perspectives claim that abstract concepts are not categorically different from concrete concepts and are embodied in the same way as concrete concepts. The difference is that the second approach requires additional properties such as social, emotional, and linguistic experiences to describe the embodiment of abstract concepts, whereas the third view considers abstract concepts to be equally grounded in the sensory-motor cortex. The results seemed to support the latter approach in that the embodiment of abstract concepts could be to some extent explained through sensory perception including interoception as one of the sensory modalities. In Experiment 1, interoception could distinctively explain many abstract concepts, as revealed in its high correlation with abstractness. Without the use of interoception, some abstract concepts would have been assumed to have no basis in the sensorimotor domain, but by incorporating interoception, it was possible to identify the sensory ground of many abstract concepts. Furthermore, Experiment 2 revealed a neurological foundation for interoception words by demonstrating modality-specific neural responses distinct from auditory words. If other sensorimotor domains, such as the body-effector related to the motor cortex (Lynott et al., 2020), are also taken into account, it may be possible to explain the grounding of more abstract concepts only through the sensory-motor system, as posited by the third approach.

On the other hand, the findings of source estimation seemed to be more in line with the second view, which asserts that the embodiment of other regions such as the ones related with emotional, social and linguistic processing is also necessary. Estimated sources of interoception words included left and right temporal pole, inferior frontal gyrus, medio-orbito prefrontal cortex, dorsolateral prefrontal cortex. Of these, the region known to be related to interoception was the inferior frontal gyrus which is connected to the insular. Activation of this area seemed to support that interoception played an important role in the embodiment of abstract concepts. However, interoception-related regions, in fact, consist of diverse areas such as the insular cortex and several regions connected to it—the anterior cingulate cortex, inferior frontal gyrus, amygdala, and somatosensory cortex (Craig, 2011; García-Cordero et al., 2017); not all of them were activated by interoception words in this experiment. In addition, the estimated regions for interoception words were more widely distributed and covered multiple regions compared to those of audio words, which were largely confined to the temporal lobes including the auditory cortex. The temporal pole, inferior frontal gyrus, medial-orbito-prefrontal cortex, dorsolateral prefrontal cortex, and amygdala were also estimated for interoception words, and those areas are known to be involved in processing emotion, social information, and language. In short, the estimated source of interoception was not limited to interoception-related areas but was more distributed over various domains related to emotion, social and linguistic experience, which was consistent with the assumption of the second approach such as WAT.

However, in light of these findings, it is difficult to conclude whether the second or third position is favored, as interoception appears to be intrinsically and inextricably correlated with emotionality and sociality. Interoception has a high correlation with emotion, and emotion has a high correlation with sociality (Villani et al., 2019). Emotion has a great influence on human cognition, behavior, and physical response (DeSteno, Gross & Kubzansky, 2013; Tyng et al., 2017), and conversely, human bodily reactions, including interoception, trigger emotions (Holzer, 2017), as in the self-attribution theory of emotion (Laird, 1974) that the expression of emotion is attributable to observation of one's own bodily reactions. Moreover, emotions are also activated by social circumstances (Padilla-Coreano, Tye, & Zelikowsky, 2022), whereas emotions influence social interactions (Van Kleef, 2009). This interdependence is reflected in the finding that *interoception*, *emotion*, *general metacognition*, and *social valence* formed one cluster ("Inner grounding and social component") when hundreds of abstract concepts were grouped together with similar properties by statistical clustering procedure in Villani et al. (2019). That they are so indistinguishable may suggest that they are also neurally interconnected. In fact, areas of the brain related to interoception, social and emotional processing substantially overlap, such as amygdala, medial-orbito prefrontal cortex, temporal pole, inferior frontal gyrus (Fossati et al., 2012; García-Cordero et al., 2017; Tso et al., 2018). The question is whether they are individual properties but only grouped together due to their correlation, or they really constitute a single group, which might be named simply as interoception or “inner grounding and social component” as suggested in Villani et al. (2019). It is difficult to answer this based on the findings of this dissertation alone, as it did not measure emotionality or social valence of words. In order to answer it, it will be necessary to classify words belonging solely to each feature and conduct a test to compare them. For example, words with high emotionality but low interoception and social valence and words with high social valence but low emotionality and interoception need to be compared.

In summary, the findings demonstrated that abstract concepts were not contrasted with concrete concepts and they were not dichotomously distinguished. Abstract concepts were also grounded in the sensorimotor domain, and interoception served as one of possible factors explaining their embodiment. However, the results of source estimation seemed to indicate the requirements of more than interoception in order to comprehensively account for the embodiment of abstract concepts. Embodiment through emotion, social and linguistic processing, as WAT claims, also needed to be taken into consideration. However, since the relationship between interoception and emotion, social and linguistic processing is highly intermingled, it is necessary to conduct further research on it.

6.3. Differences in embodiment between L1 and L2 processing

This study has the purpose of exploring L2 processing from the perspective of embodiment cognition and comparing it with L1 processing. Previous studies dealing with the embodiment of L2 processing generally showed that sensory-motor areas were involved similarly to L1 processing, but other studies found the lack of embodiment or a diminished response in L2. If embodiment is essential for language understanding, then it must be present to some extent in L2 as well. Otherwise, understanding of the language would be impossible for L2 learners. However, compared to L1, L2 has qualitative and quantitative differences in the environment of language learning and use. It is hypothesized that embodiment of language in the sensory-motor domain is formed because words and their referents or related-experiences frequently appear together, reinforcing the connection between them. However, the environment in which L2 learners come into contact with language is poorer than that of L1 speakers in terms of encountering referents in the real world. L2 learners mostly rely on explicit learning through the mediation of L1 translation equivalents rather than directly associate language with its referent. In this respect, the embodiment of language between L1 and L2 speakers may differ both quantitatively and qualitatively depending on the extent to which the environment influences its formation. If an embodiment is possibly shaped with minimal contact with the environment or a detour through L1 translations, L2 speakers will also form a generally similar embodiment despite the difference in the environment. Otherwise, they may exhibit a reduced degree of embodiment compared to L1 speakers or characteristics of L1-mediated learning.

The results of the experiments indicate that there is a comparable degree of embodiment pattern in L1 and L2 processing. In Experiment 1, L2 speakers were able to reproduce sensory experiences related to words using them, and the distribution of their perceptual strength norms was comparable to that of L1 speakers. In Experiment 2, the modality-specific neural responses were found to be similar in the early time window for both English and Korean. These findings might suggest that L2 processing is embodied in a similar manner as L1 processing, despite the inferior language environment of L2. In addition, regarding the environmental influences of L2 learning, it is interpreted that language can be grounded in sensory-motor areas even only with limited direct link with the environment or through the mediation of L1.

However, the conclusion that the embodiment of L2 is completely the same as that of L1 or that it is formed with minimal environmental interaction might be an oversimplified interpretation. This is first due to the nature of the task assigned in the experiment, which was limited to lexical processing, relatively simple in terms of language processing. Moreover, the participants in this experiment were sufficiently high-proficient learners, and the words presented to them were those with high frequency and familiarity. Therefore, the present results only pertain to L2 lexical processing that might be fully learned. To address the question of how much environmental interaction is necessary for the construction of embodiment, further research utilizing less familiar vocabulary or extending the scope to include sentences is required.

Furthermore, the findings of this study also revealed that there were several differences in neural responses between L1 and L2, as evidenced by variations in repetition enhancement, gamma band power patterns, and estimated neural generators. Specifically, the results indicated a lack of repetition enhancement in L2, an increase in gamma band power for interoception in L2, and the left temporal pole being identified as an additional neural generator for both audio and interoception in L2. These discrepancies may suggest a weaker embodiment of L2 processing or the influence of L1's intervention, or a greater dependence on linguistic experience.

In summary, this study has provided evidence that L2 processing, despite relatively limited exposure and interaction with the target language, exhibits similarities in embodiment to L1. However, the results also indicated that L2 processing may exhibit a weaker embodiment or a greater reliance on linguistic experience in comparison to L1, possibly as a result of fewer direct experiences with the target language and greater influence from the dominant language.

6.4. Theoretical implications on related issues in embodiment cognition theory and SLA

6.4.1. The connection between modality-specific regions and semantic hubs

Although the results of the current study offered support for the embodiment cognition theory of language, some of the findings also align with the amodal theory of semantic processing. The results of source estimation for lower frequencies identified the medial temporal lobe as the source, while higher frequencies revealed the temporal pole along with modality-specific regions as sources. The medial temporal lobe and temporal poles are frequently considered as semantic hubs for processing amodal semantic representations, which may appear to contradict the tenets of the embodiment cognition theory.

However, the activations of these regions are not contradictory to the point of view of embodiment cognition theory, if accompanied by modality-specific neural responses. It is compatible with most approaches with the exception of strong forms of the theory. Many embodiment cognition approaches do not assume that lexical meaning is processed solely through the involvement of modality-specific regions. Rather, they tend to employ semantic memory models that postulate a semantic or conceptual hub that serves to integrate or transmit the information of modality-specific regions. This is because there are many neurophysiological experiments demonstrating activation of language-specific areas and lesion studies related to semantic dementia (Bonner & Price, 2013; Raslau et al., 2015). Furthermore, it is not unreasonable to assume a convergence zone or a power station or a memory storage that connects various sensorimotor areas located in distributed areas of the brain (Kiefer & Pulvermüller, 2012; Patterson et al., 2007; Patterson & Ralph, 2016). For example, the hub-and-spoke model (Patterson & Ralph, 2016) posits that modality-specific conceptual knowledge is represented in modality-specific cortical regions, and that these regions exchange information with the hub in anterior temporal lobes. The difference between this point of view and the amodal theory is that it does not hold that the conceptual hub processes meaning through amodal representation but through distributed sensory-motor activation which is essential in language processing. It just emphasizes the importance of the connections between sensory-motor areas and the hub, which is called spokes (Patterson & Ralph, 2016). The results of source estimation in this study seemed to be consistent with the assumptions of this model.

In sum, modality-specific regions and so-called amodal hubs were simultaneously estimated as sources in this study, but this does not contradict the assumptions of embodiment theory. Rather, it is consistent with the theory that emphasizes the involvement of modality-specific regions and their communication in semantic processing such as the hub-and-spoke model.

6.4.2. Bilingual mental lexicon

In the examination of bilingual mental lexicon models, the utilization of perceptual strength norms might be an appropriate method. Existing psycholinguistic studies have mainly relied on various reaction time tasks such as lexical decision, semantic priming, translation equivalent recognition, word association, and semantic categorization (De Groot, 1992; Kroll, 1993). These reaction time-based tasks mostly benefit from exploring relationships between words, rather than words and their referents (Pavlenko, 2009). Definitions of lexical concepts vary, but one of the views often adapted in mental lexicon research is to regard lexical concepts as multimodal mental representations that include sensory-perceptual information stored in implicit memory (Pavlenko, 2009). From this point of view, using perceptual strengths seems to be a suitable method in exploration of bilingual mental lexicon, since it measures the relationship between words and referents and the senses surrounding them.

Among various mental lexicon models for bilingual speakers, the results of this study are most consistent with *the shared asymmetrical model* (Dong, Gui &

MacWhinney, 2005) or *the modified hierarchical model* (Pavlenko, 2009) (For more detail, see *2.2.2.2. Mental lexicon models* in *Literature review* section). A distinguishing feature of these two models is that they assume that the L1 and L2 lexicons are lexically connected and the concepts connected to the lexicons are divided into shared elements, L1-specific elements, and L2-specific elements (Dóczi, 2019). L2 lexicon is connected to all three elements, and its connection strength can vary depending on the learner's proficiency. The *shared asymmetrical model* asserts that L2 representations are closer to L2 specific concepts, whereas, since the *modified hierarchical model* assumes that L1 or shared concepts are first utilized before L2-specific elements, if they are available.

Experiment 1 demonstrated that there were shared concepts, L1-specific concepts, and L2-specific concepts. The evaluation of L2 English participants' perceptual strengths had a similar distribution to both L1 English and L1 Korean, indicating the existence of shared concepts. However, some words had a distribution more similar to L1 Korean, while others had one more comparable to L1 English. This suggests that the L2 lexicon is not only linked to shared-concepts, but also both to L1-specific concepts and L2-specific concepts. The results of the experiment also revealed that the distribution of perceptual strengths of L2 English was a slightly more similar to that of L1 Korean. This is in line with the perspective of *the modified hierarchical model*, which assumes that L1 concepts and shared concepts are employed first. Additionally, the results of this experiment also seem to prove the existence of a lexical link between L1 and L2 lexicon. In Experiment 1, the perceptual strengths of several words showed that the evaluation of L2 English participants was affected by L1 translations, and in Experiment 2, regions related with linguistic processing were activated in L2 source estimation.

6.4.3. The role of input in L2 embodiment

Exposure to input plays a crucial role in language learning. According to Emergentist approaches, language learning is accomplished through probabilistic associations in inputs by general cognitive learning mechanisms. Language learners, being repeatedly exposed to input, implicitly consider various factors such as frequency, redundancy, and saliency to establish the structure of form and meaning. This similar mechanism governs L2 learning, but in L2 learning the input is insufficient for successful learning (Ellis & Wulff, 2015). Direct exposure to L2's input is not only quantitatively less, but also qualitatively poorer than L1 input. Regarding the embodiment of L2 processing, the quantity and quality of L2 inputs are even more limited. The association between environment and language is a key input for the formation of embodiment. In the learning environment of the L2 participants of this dissertation, not only was there less direct exposure to L2, but also less interaction between L2 and the environment. It is highly likely that L2's input to embodiment is provided indirectly through text or through L1's mediation rather than direct exposure to the environment. The question is whether embodiments can be formed despite this insufficient input.

The results first demonstrated that despite these environmental constraints, L2 processing is as embodied as L1 processing. This implies that embodiment can be formed with little or indirect input. Perhaps this result is rather predictable from the standpoint of the embodiment cognition theory that embodiment is essential for the understanding of language.

Whether L1 or L2 speaker, actual language acquisition may not necessarily involve direct and continuous engagement with the environment. In many instances, the meaning of language, especially true for abstract concepts, is conveyed through a small number of contacts or in an indirect way. Therefore, the assertion of embodiment cognition theory can be established only when embodiment can be formed with a few, indirect inputs. On the other hand, the experiments also revealed that the embodiment of L2 processing was weaker than that of L1 or relying more on language experience. This showed that despite fewer and indirect input, the embodiment of L2 can be formed similarly to that of L1, but the difference in input also eventually brings about the difference in embodiment.

Nevertheless, the findings of this study can only answer that differences in input had influence on the formation of embodiment but cannot answer how much input would be needed to form an embodiment comparable to L1. Although the characteristics of the participants' L2 input were generalized based on their linguistic background, the quality and quantity of input may differ for each word. Future research should attempt to answer this by examining how the degree of embodiment varies depending on the frequency of word use and the learning environment.

6.4.4. The influence of L1 on L2 embodiment

In addition to fewer and indirect inputs, what differentiates L2 learning from L1 learning is the influence of an already established structure, that is, the learner's L1. From an emergentist point of view, the influence of L1 is present at all stages of L2 learning and use and often considered as major interference for L2 learning.

The question of this dissertation is whether and how the learner's L1 influences on L2 embodiment. The findings of this dissertations revealed the influences of the learners' L1 on L2 embodiment. In Experiment 1, L2 English participants' ratings of the perceptual strengths on several words such as *bible*, *philosophy*, *musical* and *audio* were more comparable to their L1 Korean distribution. Although they recognized words in the form of L2, the sensory perception they associated with the words was rather similar to that of their native language. According to Jiang (2000)'s model of vocabulary acquisition, it is close to the lexical association stage where L2 learners associate the meaning of the word with their L1 or the L1 lemma mediation stage the semantic information is transferred from L1 to L2 (See also Jiang, 2002). However, it is important to note that the effect of L1 on lexical processing was not dominant for the participants in this experiment, as the distribution for some words was rather similar to that of L1 English. Additionally, it is uncertain whether the effects of L1 observed in perceptual strengths were reflected in the neural responses. In Experiment 2, it was found that the left temporal pole associated with linguistic experience was activated solely in L2, regardless of the concreteness of the word. This may suggest a lexical influence of L1, this is an inference based on correlation and not an unambiguous indication of L1 influence.

Although the findings of this study revealed the influence of L1 on L2 embodiment, it is difficult to conclusively determine whether this influence was favorable or unfavorable. One potential benefit of L1 influence on L2 embodiment may be that it serves as a mechanism for compensating for a lack of input. If the embodiment of language is essential to its comprehension, then the language must be embodied in some way. The WAT posits that abstract concepts which lack concrete referents can be grounded in linguistic experiences. When this assumption is applied to L2 processing, the concepts with limited direct experience and corresponding weak associations with real-world referents can be also grounded in linguistic experiences through mediation in the learner's L1. According to the modified hierarchical mental lexicon model described above, L2 lexicon is reconstructed and developed only when L1 or shared elements are unavailable. If an existing concept is accessible, it is not necessary to create a new connection to replace it, but rather it is possible to simply use an already established structure. From this point of view, pre-existing L1 constructions may rather play a role in facilitating learning. Nevertheless, as languages do not have one-to-one correspondences, adopting already established constructs might be an obstacle to language acquisition, resulting in acquiring inaccurate forms-concept relationships. However, judging from the data of Experiment 1, in most words, L1 English participants and L2 English participants seemed to share the similar perception of the concept.

While using the established structure of L1 might be helpful in terms of learning, it could be a disadvantage for applying what has been learned. As seen earlier, the embodiment of L2 processing was relatively more dependent on linguistic experience, seemingly reflecting the use of L1 translations. Having to first access the translation of L1 instead of directly accessing the concept of L2 would be inefficient and could adversely affect L2 processing. Additionally, with the analogy of the concreteness effect, it can be assumed that a word will be processed more slowly if it is less embodied or more dependent on linguistic experience. As processing of abstract concepts that have low levels of embodiment (due to distributed grounds or heterogenous nature) or depend more on linguistic experience is often slower than concrete concepts, L2 processing might be slower for the same reasons. Indeed, the participants showed longer reaction times and lower accuracy in L2 than in L1. However, since there are many factors to consider in the comparison of L1 and L2 processing, it may be difficult to simply attribute this to embodiment relying on linguistic experience. Further research is needed to explore whether the reliance on linguistic experience in fact slows the speed of language processing.

6.5. Methodological implications

EEG source estimation techniques have been continuously employed in the previous studies, but it is still difficult to achieve high spatial resolution due to the inherent limitations of the techniques. Another major drawback is that there is no obvious statistical approach for inferential statistics (Luck, 2014). As a means of overcoming this disadvantage, the current study used four different source estimation algorithms and specified the areas as sources when consistently estimated from all four algorithms. Unless the four methods showed consistency, the estimated source was discarded from the analysis. This is comparable to measuring interrater consistency and utilizing only highly reliable sources. Although this method also does not guarantee statistical objectivity because the experimenter's subjectivity can intervene, at least this method can provide a minimal assurance of whether estimation is appropriately conducted.

In this study's source estimation analysis, two analyses were conducted by separating the frequency band into a lower frequency and higher frequency band (gamma band). Source analysis within a specific frequency band is often utilized, but rare in embodiment cognition research. The selection of the frequency bands was based on the results of the ERP and time-frequency analysis in Experiment 2 and the findings from the previous study that higher frequency bands reflect stimulus-specific neuronal representation related to cognitive processing (Pulvermüller et al., 1996). In the source estimation results, medial temporal lobes related to memory retrieval were estimated at lower frequencies, whereas cortical areas related to sensory perception were estimated as sources at higher frequencies. Activation of these two regions is compatible with the hub-and-spoke model mentioned above or the distributed network of semantic processing related to N400. However, more research is needed on the functional separation between these lower and higher frequencies, although this result is not unexpected as it is known that the cognitive function is divided according to the frequency band.

6.6. Limitations and future directions

This section discusses the limitations of this dissertation identified from the experiments and general discussion, and seeks future directions for development. Although various limitations have been recognized through the discussion of this study (e.g. the number of words evaluated in Experiment 1 was insufficient for the test to serve as the norms, or the prime task was conducted only within language but not cross-linguistically, or the results of the experiments cannot be generalized to other knowledge types such as production since it was restricted to language comprehension) the limitations will be discussed from three perspectives, methodology, material, and participants and then, the suggestion for future studies will be presented for each.

First, the first limitation is that many of the conclusions of this study depended on the findings of source estimation analysis. As described above, various methods were introduced to guarantee minimum validity, but it was not possible to fully overcome the limitations of the low spatial resolution of EEG source estimation. Although, due to the complexity of the brain's cognitive functions, no single area represents the entirety of a specific function, sometimes adjacent regions reflect completely different functions. Consequently, a minor estimation inaccuracy might result in a completely different interpretation. In this paper, not only ERP and time-frequency analysis, but also the reaction times of the lexical decision task and perceptual strengths were discussed altogether, but in embodiment cognition theory, source estimation inevitably plays a decisive role in interpreting the results. In the future studies, it will be required to develop more sophisticated and reliable EEG source estimation analysis techniques such as state-of-the-art techniques using machine learning or deep learning, as the use of EEG for assessing early processing will continue to be a necessity in this research field.

Secondly, the sensory modalities dealt with in Experiment 2 were restricted only to a subset of the numerous senses. In this research, audio and interoception were selected and compared as the senses most related to abstract concepts, and discussions were developed based on differences in modality-specific responses between them. Different groups of senses or features may be represented in modality-specific brain areas in different ways and may have different characteristics in terms of abstract concepts. Therefore, for a more comprehensive discussion, future research needs to compare a wider variety of senses and features. In particular, this study explored the embodiment of abstract concepts through interoception, but the results of this study implied the possibility that other senses or properties than interoception seemed to be involved in the embodiment of abstract concepts. Therefore, other factors regarded as related to abstract concepts need to be explored. For instances, it is possible to investigate what implications emotional and social experience have as an embodiment and how they can be distinguished from interoception.

Finally, it is necessary to expand and diversify the group of participants. In Experiment 2, there was no control group of native English speakers. Obviously, the within-design experiments have the advantage of being able to perform comparisons within participants who have the same neuroanatomical structure and cultural, environmental influences. As shown in Experiment 1, L2 English participants seemed to have a perceptual experience more similar to that of L1 Korean participants. Therefore, comparing L1 Korean and L2 English of Korean native speakers was advantageous in exploiting this similarity in perceptual experiences. However, the major disadvantage of the within-design is that differences in the results might be caused by differences in language. Since Korean and English do not exactly correspond to each other, despite careful material selection, the present findings may be due to differences between languages rather than differences between L1 and L2. For example, this experiment found that the left temporal pole was characteristically activated in L2 processing. However, if its activation was equally observed for L1 native speakers, it would be not a characteristic of L2 processing, but of English processing or selected English words.

In addition, L2 participants in this study represented only a small subgroup of the numerous types of L2 learners. L2 participants in this study were a certain type of L2 learners who learned and used L2 in foreign settings. The characteristics and language environment of L2 learners described in this dissertation focused on this type of participants, and the results were interpreted accordingly. As the data only reflected the embodiment of L2 learners with certain characteristics, it would be difficult to generalize the findings of this study to all L2 learners. In the future studies, it is necessary to compare L2 participants with more diverse individual variables, in order to explore which characteristics of L2 processing were responsible for weakened embodiment or the involvement of linguistic experience observed in Experiment 2. For example, participants with higher proficient or more experiences in language use may show more L1-like embodiment formation.

In this study, comparison analyses on various L2 subgroups were not conducted due to the relatively limited number of participants. Instead, four participants with disparate language use experience were removed from the analysis. Since four participants were too few to be statistically meaningful, it was not possible to separately analyze them as a different group. However, if possible, it is important to conduct further analysis using a technique that can reliably infer the statistical differences even with a small number of participants in the future studies. For example, with modeling based on deep learning techniques, analysis can be conducted with only a small number of trials (Bagchi & Bathula, 2022; Zang et al., 2021). In future research, if analysis is conducted by dividing L2 participants based on proficiency or language usage experience, richer discussions on L2 processing and embodiment will be able to develop.

Chapter 7. Conclusion

In conclusion, from the embodiment cognition theory of language processing, this dissertation explores whether the activation of the sensorimotor areas of brain is essential for language processing, whether abstract concepts are also grounded in the sensorimotor areas, and whether L2 processing activates the sensorimotor areas in the same manner as L1 processing. To this end, perceptual strength norms of L1 Korean, L1 English, and L2 English were constructed to examine sensory experiences associated with words, and EEG experiments were conducted to investigate whether modality-specific neural responses were elicited from interoception and audio. The results of the experiments demonstrated that the sensory motor areas were engaged from the early time windows of language processing, that the sensorimotor areas related to interoception were activated along with several other areas in the processing of abstract concepts, and that L2 processing was also embodied similarly to L1 although its embodiment was weaker and more dependent on linguistic experience. The findings of this dissertation provided evidence of supporting the embodiment cognition theory of language processing and suggested the necessity to pay attention to the differences in embodiment cognition induced by the learning environment in terms of SLA. In addition, methodologically, this dissertation introduced a new approach that might enhance the validity of the EEG source estimation technique. Along with these contributions, several limitations were also identified. It is necessary to investigate features other than interoception, such as emotion, social and language experience, in order to fully understand the embodiment of abstract concepts. In addition, it is also important to compare the L2 learner groups with different individual characteristics to the L1 native control group, in order to explore the influence of the language learning environment on embodiment and L2 processing. Future studies will not only contribute to a deeper understanding of the embodied cognition theory and SLA theories, but also of human language processing in general.

References

Ablin, P., Cardoso, J. F., & Gramfort, A. (2018). Faster independent component analysis by preconditioning with Hessian approximations. *IEEE Transactions on Signal Processing, 66(15)*, 4040-4049.

Addante, R. J., Watrous, A. J., Yonelinas, A. P., Ekstrom, A. D., & Ranganath, C. (2011). Prestimulus theta activity predicts source memory retrieval. *Proceedings of the National Academy of Sciences, 108(26),* 10702-10707.

Adelman, J. S., Johnson, R. L., McCormick, S. F., McKague, M., Kinoshita, S., Bowers, J. S., ... & Davis, C. J. (2014). A behavioral database for masked form priming. *Behavior research methods, 46(4),* 1052-1067.

Ahlberg, D. K., Bischoff, H., Kaup, B., Bryant, D., & Strozyk, J. V. (2018). Grounded cognition: Comparing language × space interactions in first language and second language. *Applied Psycholinguistics*, *39*(2), 437-459.

Alcón, E. (2007). Incidental focus on form, noticing and vocabulary learning in the EFL classroom. *International Journal of English Studies*, *7*(2), 41-60.

Altarriba, J., Bauer, L. M., & Benvenuto, C. (1999). Concreteness, context availability, and imageability ratings and word associations for abstract, concrete, and emotion words. *Behavior Research Methods, Instruments, & Computers, 31(4),* 578-602.

Anderer, P., Semlitsch, H. V., & Saletu, B. (1996). Multichannel auditory event-related brain potentials: effects of normal aging on the scalp distribution of N1, P2, N2 and P300 latencies and amplitudes.

Electroencephalography and clinical Neurophysiology, 99(5), 458-472.

Andrews, M., Frank, S., & Vigliocco, G. (2014). Reconciling embodied and distributional accounts of meaning in language. *Topics in cognitive science, 6(3),* 359-370.

Ardila, A., Bernal, B., & Rosselli, M. (2014). The elusive role of the left temporal pole (BA38) in language: a preliminary meta-analytic connectivity study. *International Journal of Brain Science, 2014*.

Bagchi, S., & Bathula, D. R. (2022). EEG-ConvTransformer for single-trial EEG-based visual stimulus classification. *Pattern Recognition, 129*, 108757.

Bak, T. H., O'Donovan, D. G., Xuereb, J. H., Boniface, S., & Hodges, J. R. (2001). Selective impairment of verb processing associated with pathological changes in Brodmann areas 44 and 45 in the motor neurone disease–dementia–aphasia syndrome. *Brain, 124*(1), 103-120.

Banich, M. T., & Compton, R. J. (2018). *Cognitive neuroscience*. Cambridge University Press.

Barber, H. A., Otten, L. J., Kousta, S. T., & Vigliocco, G. (2013). Concreteness in word processing: ERP and behavioral effects in a lexical decision task. *Brain and language, 125*(1), 47-53.

Barsalou, L. W., Santos, A., Simmons, W. K., & Wilson, C. D. (2008). Language and simulation in conceptual processing. *Symbols, embodiment, and meaning*, 245-283.

Barsalou, L. W., & Wiemer-Hastings, K. (2005). Situating abstract concepts. *Grounding cognition: The role of perception and action in memory, language, and thought*, 129-163.

Başar, E., Başar-Eroglu, C., Karakaş, S., & Schürmann, M. (2001). Gamma, alpha, delta, and theta oscillations govern cognitive processes. *International journal of psychophysiology, 39(2-3)*, 241-248.

Başar-Eroglu, C., Strüber, D., Kruse, P., Başar, E., & Stadler, M. (1996). Frontal gamma-band enhancement during multistable visual perception. *International Journal of Psychophysiology*, 24(1-2), 113-125.

Bastiaansen, M., & Hagoort, P. (2006). Oscillatory neuronal dynamics during language comprehension. *Progress in brain research, 159*, 179-196.

Bastiaansen, M., & Hagoort, P. (2015). Frequency-based segregation of syntactic and semantic unification during online sentence level language comprehension. *Journal of cognitive neuroscience*, *27(11)*, 2095-2107.

Bastiaansen, M. C., Van Der Linden, M., Ter Keurs, M., Dijkstra, T., & Hagoort, P. (2005). Theta responses are involved in lexical—Semantic retrieval during language processing. *Journal of cognitive neuroscience, 17(3),* 530-541.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software, 67(1)*, 1-48.

Bayley, P. J., & Squire, L. R. (2005). Failure to acquire new semantic knowledge in patients with large medial temporal lobe lesions. *Hippocampus, 15*(2), 273-280.

Beauprez, S. A., & Bidet-Ildei, C. (2019). Perceiving a biological human movement facilitates action verb processing. *Current Psychology, 38(5),* 1355-1359.

Beauprez, S. A., Blandin, Y., Almecija, Y., & Bidet-Ildei, C. (2020). Physical and observational practices of unusual actions prime action verb processing. *Brain and cognition, 138*, 103630.

Bechtold, L., Bellebaum, C., Egan, S., Tettamanti, M., & Ghio, M. (2019). The role of experience for abstract concepts: Expertise modulates the electrophysiological correlates of mathematical word processing. *Brain and language, 188*, 1-10.

Benjamini, Y. (2010). Simultaneous and selective inference: Current successes and future challenges. *Biometrical Journal*, *52*(6), 708-721.

Bergelson, E., & Swingley, D. (2013). The acquisition of abstract words by young infants. *Cognition, 127(3)*, 391-397.

Bergen, B., Lau, T. T. C., Narayan, S., Stojanovic, D., & Wheeler, K. (2010). Body part representations in verbal semantics. *Memory & cognition, 38(7),* 969-981.

Bidet-Ildei, C., Gimenes, M., Toussaint, L., Almecija, Y., & Badets, A. (2017). Sentence plausibility influences the link between action words and the perception of biological human movements. *Psychological research, 81(4)*, 806-813.

Binder, J. R., Westbury, C. F., McKiernan, K. A., Possing, E. T., & Medler, D. A. (2005). Distinct brain systems for processing concrete and abstract concepts. *Journal of cognitive neuroscience, 17(6),* 905-917.

Birba, A., Beltrán, D., Caro, M. M., Trevisan, P., Kogan, B., Sedeño, L., ... & García, A. M. (2020). Motor-system dynamics during naturalistic reading of action narratives in first and second language. *Neuroimage, 216,* 116820.

Birot, G., Spinelli, L., Vulliémoz, S., Mégevand, P., Brunet, D., Seeck, M., & Michel, C. M. (2014). Head model and electrical source imaging: a study of 38 epileptic patients. *NeuroImage: Clinical*, 5, 77-83.

Blank, S. C., Scott, S. K., Murphy, K., Warburton, E., & Wise, R. J. (2002). Speech production: Wernicke, Broca and beyond. *Brain, 125*(8), 1829-1838.

Blomberg, F., Roll, M., Frid, J., Lindgren, M., & Horne, M. (2020). The role of affective meaning, semantic associates, and orthographic neighbours in modulating the N400 in single words. *The Mental Lexicon, 15*(2), 161-188.

Bock, K., Levelt, W., & Gernsbacher, M. A. (2002). Language production: Grammatical encoding. *Psycholinguistics: Critical concepts in psychology, 5,* 405-452.

Bonner, M. F., & Price, A. R. (2013). Where is the anterior temporal lobe and what does it do?. *Journal of Neuroscience, 33*(10), 4213-4215.

Borghi, A. M. (2020). A future of words: Language and the challenge of abstract concepts. *Journal of Cognition, 3(1)*

Borghi, A. M., Barca, L., Binkofski, F., Castelfranchi, C., Pezzulo, G., & Tummolini, L. (2019). Words as social tools: Language, sociality and inner grounding in abstract concepts. *Physics of life reviews, 29,* 120-153.

Borghi, A. M., Binkofski, F., Castelfranchi, C., Cimatti, F., Scorolli, C., & Tummolini, L. (2017). The challenge of abstract concepts. *Psychological Bulletin, 143(3),* 263.

Borghi, A. M., & Zarcone, E. (2016). Grounding abstractness: abstract concepts and the activation of the mouth. *Frontiers in psychology, 7,* 1498.

Brooks, R., & Meltzoff, A. N. (2005). The development of gaze following and its relation to language. *Developmental science, 8(6),* 535-543.

Brown, R., Waring, R., & Donkaewbua, S. (2008). Incidental vocabulary acquisition from reading, reading-while-listening, and listening to stories. *Reading in a Foreign Language, 20*(2), 136–63.

Brysbaert, M., Warriner, A. B., & Kuperman, V. (2014). Concreteness ratings for 40 thousand generally known English word lemmas. *Behavior research methods*, 46(3), 904-911.

Buccino, G., Colagè, I., Silipo, F., & D'Ambrosio, P. (2019). The concreteness of abstract language: an ancient issue and a new perspective. *Brain Structure and Function, 224(4),* 1385-1401.

Buccino, G., Marino, B. F., Bulgarelli, C., & Mezzadri, M. (2017). Fluent speakers of a second language process graspable nouns expressed in L2 like in their native language. *Frontiers in Psychology, 8,* 1306.

Buresh, J. S., & Woodward, A. L. (2007). Infants track action goals within and across agents. *Cognition, 104(2),* 287-314.

Buzsaki, G., & Draguhn, A. (2004). Neuronal oscillations in cortical networks. *Science, 304(5679),* 1926-1929.

Burgess, C., & Lund, K. (1997). Modeling parsing constraints with high-dimensional contextspace. *Language and Cognitive Processes, 12,* 177–210.

Canolty, R. T., & Knight, R. T. (2010). The functional role of cross-frequency coupling. *Trends in cognitive sciences, 14(11),* 506-515.

Carr, T. H., & Dagenbach, D. (1990). Semantic priming and repetition priming from masked words: evidence for a center-surround attentional mechanism in perceptual recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 16(2), 341.

Carrillo-de-la-Peña, M. T., & Cadaveira, F. (2000). The effect of motivational instructions on P300 amplitude. *Neurophysiologie Clinique/Clinical Neurophysiology, 30*(4), 232-239.

Cervetto, S., Díaz-Rivera, M., Petroni, A., Birba, A., Caro, M. M., Sedeño, L., Ibáñez, I., & García, A. M. (2021). The Neural Blending of Words and Movement: Event-Related Potential Signatures of Semantic and Action Processes during Motor–Language Coupling. *Journal of Cognitive Neuroscience, 33(8)*, 1413-1427.

Catricalà, E., Della Rosa, P. A., Plebani, V., Vigliocco, G., & Cappa, S. F. (2014). Abstract and concrete categories? Evidences from neurodegenerative diseases. *Neuropsychologia*, *64*, 271-281.

Chambers, C. G., Tanenhaus, M. K., & Magnuson, J. S. (2004). Actions and affordances in syntactic ambiguity resolution. *Journal of experimental psychology: Learning, memory, and cognition, 30(3)*, 687.

Chedid, G., Brambati, S. M., Bedetti, C., Rey, A. E., Wilson, M. A., & Vallet, G. T. (2019). Visual and auditory perceptual strength norms for 3,596 French nouns and their relationship with other psycholinguistic variables. *Behavior research methods*, *51*(5), 2094-2105.

Chen, I. H., Zhao, Q., Long, Y., Lu, Q., & Huang, C. R. (2019). Mandarin Chinese modality exclusivity norms. *PloS one*, *14*(2), e0211336.

Chomsky, N. (1959). A review of BF Skinner's *Verbal Behavior*. *Language*, *35*(1), 26–58.

Citron, F. M. (2012). Neural correlates of written emotion word processing: a review of recent electrophysiological and hemodynamic neuroimaging studies. *Brain and language, 122*(3), 211-226.

Clark, J. M., & Paivio, A. (1987). A dual coding perspective on encoding processes.

In Imagery and related mnemonic processes (pp. 5-33). Springer, New York, NY.

Cohen, M. X. (2014). *Analyzing neural time series data: theory and practice*. MIT press.

Cohen, N. J., Poldrack, R. A., & Eichenbaum, H. (1997). Memory for items and memory for relations in the procedural/declarative memory framework. *Memory, 5(1-2)*, 131-178.

Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological review, 82*(6), 407.

Collins, A. M., & Quillian, M. R. (1969). Retrieval time from semantic memory. *Journal of verbal learning and verbal behavior, 8*(2), 240-247.

Conca, F., Borsa, V. M., Cappa, S. F., & Catricalà, E. (2021). The multidimensionality of abstract concepts: A systematic review. *Neuroscience & Biobehavioral Reviews, 127*, 474-491.

Connell, L., & Lynott, D. (2012). Strength of perceptual experience predicts word processing performance better than concreteness or imageability. *Cognition, 125(3)*, 452-465.

Connell, L., & Lynott, D. (2016). Do we know what we're simulating? Information loss on transferring unconscious perceptual simulation to conscious imagery. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *42*(8), 1218.

Connell, L., Lynott, D., & Banks, B. (2018). Interoception: the forgotten modality in perceptual grounding of abstract and concrete concepts. *Philosophical Transactions of the Royal Society B: Biological Sciences, 373(1752)*, 20170143.

Cotelli, M., Borroni, B., Manenti, R., Alberici, A., Calabria, M., Agosti, C., ... & Cappa, S. F. (2006). Action and object naming in frontotemporal dementia, progressive supranuclear palsy, and corticobasal degeneration. *Neuropsychology*, *20*(5), 558.

Critchley, H. D., & Harrison, N. A. (2013). Visceral influences on brain and behavior. *Neuron*, *77(4)*, 624-638.

Craig, A. D. (2003). Interoception: the sense of the physiological condition of the body. *Current opinion in neurobiology, 13(4)*, 500-505.

Craig, A. D. (2011). Significance of the insula for the evolution of human awareness of feelings from the body. *Annals of the new York Academy of Sciences, 1225(1)*, 72-82.

Crossley, S. A., Subtirelu, N., & Salsbury, T. (2013). Frequency effects or context effects in second language word learning: What predicts early lexical production?. *Studies in Second Language Acquisition*, *35*(4), 727-755.

Crowley, K. E., & Colrain, I. M. (2004). A review of the evidence for P2 being an independent component process: age, sleep and modality. *Clinical neurophysiology*, *115*(4), 732-744.

Curtis, C. E., & D'Esposito, M. (2003). Persistent activity in the prefrontal cortex during working memory. *Trends in cognitive sciences*, *7(9)*, 415-423.

Dale, A. M., Liu, A. K., Fischl, B. R., Buckner, R. L., Belliveau, J. W., Lewine, J. D., & Halgren, E. (2000). Dynamic statistical parametric mapping: combining fMRI and MEG for high-resolution imaging of cortical activity. *neuron*, *26*(1), 55-67.

Dalla Volta, R., Avanzini, P., De Marco, D., Gentilucci, M., & Fabbri-Destro, M. (2018). From meaning to categorization: the hierarchical recruitment of brain circuits selective for action verbs. *Cortex, 100*, 95-110.

Dalla Volta, R., Fabbri-Destro, M., Gentilucci, M., & Avanzini, P. (2014). Spatiotemporal dynamics during processing of abstract and concrete verbs: an ERP study. *Neuropsychologia, 61*, 163-174.

David, O., Kilner, J. M., & Friston, K. J. (2006). Mechanisms of evoked and induced responses in MEG/EEG. *Neuroimage, 31(4),* 1580-1591.

Davidson, D. J., & Indefrey, P. (2007). An inverse relation between event-related and time–frequency violation responses in sentence processing. *Brain research, 1158,* 81-92.

DeKeyser, R. (2015). Skill Acquisition Theory. In *Theories in Second Language Acquisition* (pp. 106-124). Routledge.

De Grauwe, S., Willems, R. M., Rueschemeyer, S. A., Lemhöfer, K., & Schriefers, H. (2014). Embodied language in first-and second-language speakers: Neural correlates of processing motor verbs. *Neuropsychologia, 56*, 334-349.

De Groot, A. M. (1992). Bilingual lexical representation: A closer look at conceptual representations. In *Advances in psychology* (Vol. 94, pp. 389-412). North-Holland.

Dehaene, S., Dupoux, E., Mehler, J., Cohen, L., Paulesu, E., Perani, D., ... & Le Bihan, D. (1997). Anatomical variability in the cortical representation of first and second language. *Neuroreport, 8*(17), 3809-3815.

Desai, R. H., Binder, J. R., Conant, L. L., Mano, Q. R., & Seidenberg, M. S. (2011). The neural career of sensory-motor metaphors. *Journal of cognitive neuroscience, 23*(9), 2376-2386.

Dóczi, B. (2019). An overview of conceptual models and theories of lexical representation in the mental lexicon. In The *Routledge handbook of vocabulary studies* (pp. 46-65). Routledge.

Dong, Y., Gui, S., & MacWhinney, B. (2005). Shared and separate meanings in the bilingual mental lexicon. *Bilingualism: Language and Cognition, 8*(3), 221-238.

Dove, G. (2009). Beyond perceptual symbols: A call for representational pluralism. *Cognition, 110(3)*, 412-431.

Dove, G. (2011). On the need for embodied and dis-embodied cognition. *Frontiers in Psychology*, 1, 242.

Dove, G., Barca, L., Tummolini, L., & Borghi, A. M. (2020). Words have a weight: Language as a source of inner grounding and flexibility in abstract concepts. *Psychological Research*, 1-17.

Dreyer, F. R., Frey, D., Arana, S., Saldern, S. V., Picht, T., Vajkoczy, P., & Pulvermüller, F. (2015). Is the motor system necessary for processing action and abstract emotion words? Evidence from focal brain lesions. *Frontiers in Psychology*, *6*, 1661.

Dreyer, F. R., & Pulvermüller, F. (2018). Abstract semantics in the motor system?–An event-related fMRI study on passive reading of semantic word categories carrying abstract emotional and mental meaning. *Cortex*, *100*, 52-70.

Dudschig, C., De la Vega, I., & Kaup, B. (2014). Embodiment and second-language: Automatic activation of motor responses during processing spatially associated L2 words and emotion L2 words in a vertical Stroop paradigm. *Brain and Language, 132*, 14-21.

Đurđević, D. F., Popović Stijačić, M., & Karapandžić, J. (2016). A quest for sources of perceptual richness: Several candidates. *Studies in language and mind*, 187-238.

Dunn, B. R., Dunn, D. A., Languis, M., & Andrews, D. (1998). The relation of ERP components to complex memory processing. *Brain and cognition, 36*(3), 355-376.

Eddy, M., Schmid, A., & Holcomb, P. J. (2006). Masked repetition priming and event-related brain potentials: A new approach for tracking the time-course of object perception. *Psychophysiology*, *43*(6), 564-568.

Ekstrom, A. D., Caplan, J. B., Ho, E., Shattuck, K., Fried, I., & Kahana, M. J. (2005). Human hippocampal theta activity during virtual navigation. *Hippocampus, 15(7)*, 881-889.

Elgort, I., & Warren, P. (2014). L2 vocabulary learning from reading: Explicit and tacit lexical knowledge and the role of learner and item variables. *Language Learning, 64*(2), 365-414.

Ellis, N. C. (1998). Emergentism, connectionism and language learning. *Language learning*, *48*(4), 631-664.

Ellis, N. C. (2006). Selective attention and transfer phenomena in L2 acquisition: Contingency, cue competition, salience, interference, overshadowing, blocking, and perceptual learning. *Applied linguistics*, *27*(2), 164-194.

Ellis, N. C. (2012). Frequency-based accounts of second language acquisition. *The Routledge handbook of second language acquisition*, 193-210.

Ellis, N. C., & Wulff, S. (2015). Usage-Based Approaches to SLA. In *Theories in Second Language Acquisition* (pp. 87-105). Routledge.

Eulitz, C., Maess, B., Pantev, C., Friederici, A. D., Feige, B., & Elbert, T. (1996). Oscillatory neuromagnetic activity induced by language and nonlanguage stimuli. *Cognitive brain research, 4*(2), 121-132.

Federmeier, K. D. (2022). Connecting and considering: Electrophysiology provides insights into comprehension. *Psychophysiology, 59*(1), e13940.

Fiebach, C. J., & Friederici, A. D. (2004). Processing concrete words: fMRI evidence against a specific right-hemisphere involvement. *Neuropsychologia, 42(1)*, 62-70.

Fini, C., Zannino, G. D., Orsoni, M., Carlesimo, G. A., Benassi, M., & Borghi, A. M. (2021). Articulatory suppression delays processing of abstract words: The role of inner speech. *Quarterly Journal of Experimental Psychology*. https://doi.org/10.1177/ 17470218211053623

Finnigan, S., Humphreys, M. S., Dennis, S., & Geffen, G. (2002). ERP 'old/new'effects: memory strength and decisional factor (s). *Neuropsychologia, 40*(13), 2288-2304.

Fischl, B. (2012). FreeSurfer. *Neuroimage, 62(2)*, 774-781.

Fodor, J. A. (1983). *The modularity of mind*. MIT press.

Forster, K. I., Mohan, K., Hector, J., Kinoshita, S., & Lupker, S. J. (2003). The mechanics of masked priming. *Masked priming: The state of the art*, 3-37.

Foxe, J. J., & Snyder, A. C. (2011). The role of alpha-band brain oscillations as a sensory suppression mechanism during selective attention. *Frontiers in psychology, 2,* 154.

French, R. M., & Jacquet, M. (2004). Understanding bilingual memory: Models and data. *Trends in cognitive sciences, 8*(2), 87-93.

Friederici, A. D. (2017). *Language in our brain: The origins of a uniquely human capacity.* MIT Press.

Friedman, D., & Johnson Jr, R. (2000). Event-related potential (ERP) studies of memory encoding and retrieval: A selective review. *Microscopy research and technique*, *51*(1), 6-28.

Fries, P. (2009). Neuronal gamma-band synchronization as a fundamental process in cortical computation. *Annual review of neuroscience, 32*(1), 209-224.

Funnell, M. G., Corballis, P. M., & Gazzaniga, M. S. (2001). Hemispheric processing asymmetries: implications for memory. *Brain and Cognition, 46(1-2)*, 135-139.

Gage, N. M. & Baars, B. (2018). *Fundamentals of cognitive neuroscience: a beginner's guide.* Academic Press.

García-Cordero, I., Esteves, S., Mikulan, E. P., Hesse, E., Baglivo, F. H., Silva, W., ... & Sedeño, L. (2017). Attention, in and out: scalp-level and intracranial EEG correlates of interoception and exteroception. *Frontiers in neuroscience, 11*, 411.

Garcia-Gamez, A. B., & Macizo, P. (2019). Learning nouns and verbs in a foreign language: The role of gestures. *Applied Psycholinguistics, 40(2),* 473-507.

Gass, S. M., Behney, J., & Plonsky, L. (2020). *Second language acquisition: An introductory course.* Routledge.

Gasser, T., & Molinari, L. (1996). The analysis of the EEG. *Statistical Methods in Medical Research, 5*(1), 67-99.

Ghio, M., Vaghi, M. M. S., & Tettamanti, M. (2013). Fine-grained semantic categorization across the abstract and concrete domains. *PloS one, 8(6),* e67090.

Glenberg, A. M., & Kaschak, M. P. (2002). Grounding language in action. *Psychonomic bulletin & review, 9*(3), 558-565.

Glenberg, A. M., Sato, M., & Cattaneo, L. (2008). Use-induced motor plasticity affects the processing of abstract and concrete language. *Current Biology*, *18*(7), R290-R291.

Glenberg, A. M., Sato, M., Cattaneo, L., Riggio, L., Palumbo, D., & Buccino, G. (2008). Processing abstract language modulates motor system activity. *Quarterly Journal of Experimental Psychology, 61(6),* 905-919.

Gogolla, N. (2017). The insular cortex. *Current Biology, 27*(12), R580-R586.

Gola, M., Kamiński, J., Brzezicka, A., & Wróbel, A. (2012). Beta band oscillations as a correlate of alertness—Changes in aging. *International Journal of Psychophysiology, 85(1),* 62-67.

Gramfort, A., Luessi, M., Larson, E., Engemann, D. A., Strohmeier, D., Brodbeck, C., ... & Hämäläinen, M. (2013). MEG and EEG data analysis with MNE-Python. *Frontiers in neuroscience, 267.*

Gregg, K. R. (2003). The state of emergentism in second language acquisition. *Second Language Research*, *19*(2), 95-128.

Grill-Spector, K., Henson, R., & Martin, A. (2006). Repetition and the brain: neural models of stimulus-specific effects. *Trends in cognitive sciences*, *10*(1), 14-23.

Grossman, M., Koenig, P., DeVita, C., Glosser, G., Alsop, D., Detre, J., & Gee, J. (2002). The neural basis for category-specific knowledge: an fMRI study. *Neuroimage, 15(4),* 936-948.

Gruber, T., Müller, M. M., Keil, A., & Elbert, T. (1999). Selective visual-spatial attention alters induced gamma band responses in the human EEG. *Clinical neurophysiology, 110*(12), 2074-2085.

Günther, F., Petilli, M. A., Vergallito, A., & Marelli, M. (2020). Images of the unseen: Extrapolating visual representations for abstract and concrete words in a data-driven computational model. *Psychological Research*, 1-21.

Hald, L. A., Bastiaansen, M. C., & Hagoort, P. (2006). EEG theta and gamma responses to semantic violations in online sentence processing. Brain and language, 96(1), 90-105.

Hall, J. E. (2021). States of Brain Activity-Sleep, Brain Waves, Epilepsy, Psychoses, and Dementia. *Guyton and Hall Textbook of Medical Physiology*, 763-766.

Hämäläinen, M. S., & Ilmoniemi, R. J. (1994). Interpreting magnetic fields of the brain: minimum norm estimates. *Medical & biological engineering & computing*, *32*(1), 35-42.

Han, S., Hu, J., Li, W., Zhao, S., Chen, M., Xu, P., & Luo, Y. (2021). From structure to concepts: The two stages of facial expression recognition. *Neuropsychologia*, *150*, 107700.

Harnad, S. (1990). The symbol grounding problem. *Physica D: Nonlinear Phenomena*, *42*(1-3), 335-346.

Hart, J., & Gordon, B. (1992). Neural subsystems for object knowledge. *Nature*, *359*(6390), 60-64.

Harpaintner, M., Sim, E. J., Trumpp, N. M., Ulrich, M., & Kiefer, M. (2020). The grounding of abstract concepts in the motor and visual system: An fMRI study. *Cortex*, *124*, 1-22.

Harpaintner, M., Trumpp, N. M., & Kiefer, M. (2020). Time course of brain activity during the processing of motor-and vision-related abstract concepts: Flexibility and task dependency. *Psychological Research*, 1-23.

Hauser, M. D., Chomsky, N., & Fitch, W. T. (2002). The faculty of language: what is it, who has it, and how did it evolve?. Science, 298(5598), 1569-1579.

Hauk, O. (2004). Keep it simple: a case for using classical minimum norm estimation in the analysis of EEG and MEG data. *Neuroimage, 21(4),* 1612-1621.

Hauk, O., Johnsrude, I., & Pulvermüller, F. (2004). Somatotopic representation of action words in human motor and premotor cortex. *Neuron, 41(2),* 301-307.

Hauk, O., & Pulvermüller, F. (2004). Neurophysiological distinction of action words in the fronto-central cortex. *Human brain mapping, 21(3),* 191-201.

Hebb, D. O. (2005). *The organization of behavior: A neuropsychological theory.* Psychology Press.

Herbert, C., Kissler, J., Junghöfer, M., Peyk, P., & Rockstroh, B. (2006). Processing of emotional adjectives: Evidence from startle EMG and ERPs. *Psychophysiology, 43*(2), 197-206.

Herlin, B., Navarro, V., & Dupont, S. (2021). The temporal pole: From anatomy to function—A literature appraisal. *Journal of Chemical Neuroanatomy, 113*, 101925.

Herrmann, C. S., Lenz, D., Junge, S., Busch, N. A., & Maess, B. (2004). Memory-matches evoke human gamma-responses. *BMC neuroscience, 5*(1), 1-8.

Herrmann, C. S., Munk, M. H., & Engel, A. K. (2004). Cognitive functions of gamma-band activity: memory match and utilization. *Trends in cognitive sciences, 8*(8), 347-355.

Hillyard, S. A., & Anllo-Vento, L. (1998). Event-related brain potentials in the study of visual selective attention. *Proceedings of the National Academy of Sciences, 95*(3), 781-787.

Hillyard, S. A., Teder-Sälejärvi, W. A., & Münte, T. F. (1998). Temporal dynamics of early perceptual processing. *Current opinion in neurobiology, 8*(2), 202-210.

Hinojosa, J. A., Carretié, L., Valcárcel, M. A., Méndez-Bértolo, C., & Pozo, M. A. (2009). Electrophysiological differences in the processing of affective information in words and pictures. *Cognitive, Affective, & Behavioral Neuroscience, 9*(2), 173-189.

Hipp, J. F., Engel, A. K., & Siegel, M. (2011). Oscillatory synchronization in large-scale cortical networks predicts perception. *Neuron, 69*(2), 387-396.

Hoenig, K., Sim, E. J., Bochev, V., Herrnberger, B., & Kiefer, M. (2008). Conceptual flexibility in the human brain: dynamic recruitment of semantic maps from visual, motor, and motion-related areas. *Journal of Cognitive Neuroscience, 20*(10), 1799-1814.

Hoffman, P., Jefferies, E., & Ralph, M. A. L. (2010). Ventrolateral prefrontal cortex plays an executive regulation role in comprehension of abstract words: convergent neuropsychological and repetitive TMS evidence. *Journal of Neuroscience, 30(46)*, 15450-15456.

Holcomb, P. J., & Grainger, J. (2007). Exploring the temporal dynamics of visual word recognition in the masked repetition priming paradigm using event-related potentials. *Brain research, 1180*, 39-58.

Holcomb, P. J., Kounios, J., Anderson, J. E., & West, W. C. (1999). Dual-coding, context-availability, and concreteness effects in sentence comprehension: an electrophysiological investigation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25(3),* 721.

Holcomb, P. J., Reder, L., Misra, M., & Grainger, J. (2005). Masked priming: An event-related brain potential study of repetition and semantic effects. *Cognitive Brain Research, 24(1),* 155-172.

Hsu, C. H., Tsai, J. L., Lee, C. Y., & Tzeng, O. J. L. (2009). Orthographic combinability and phonological consistency effects in reading Chinese phonograms: an event-related potential study. *Brain and Language, 108*(1), 56-66.

Hulstijn, J. H. (2005). Theoretical and empirical issues in the study of implicit and explicit second-language learning: Introduction. *Studies in second language acquisition, 27(2),* 129-140.

Ibáñez, A., Manes, F., Escobar, J., Trujillo, N., Andreucci, P., & Hurtado, E. (2010). Gesture influences the processing of figurative language in non-native speakers: ERP evidence. *Neuroscience Letters, 471(1),* 48-52.

Ihara, A., Hirata, M., Sakihara, K., Izumi, H., Takahashi, Y., Kono, K., ... & Yorifuji, S. (2003). Gamma-band desynchronization in language areas reflects syntactic process of words. *Neuroscience Letters, 339*(2), 135-138.

James, C. T. (1975). The role of semantic information in lexical decisions. *Journal of Experimental Psychology: Human Perception and Performance, 1(2),* 130.

Jarema, G., & Libben, G. (2007). Introduction: Matters of Definition and Core Perspectives. In *The Mental Lexicon* (pp. 1-6). Brill.

Jas, M., Engemann, D. A., Bekhti, Y., Raimondo, F., & Gramfort, A. (2017). Autoreject: Automated artifact rejection for MEG and EEG data. *NeuroImage, 159*, 417-429.

Jatoi, M. A., & Kamel, N. (2017). *Brain source localization using EEG signal analysis.* CRC Press.

Jiang, N. (2000). Lexical representation and development in a second language. *Applied linguistics, 21*(1), 47-77.

Jiang, N. (2002). Form–meaning mapping in vocabulary acquisition in a second language. *Studies in Second Language Acquisition, 24*(4), 617-637.

Jiang, N. (2004). Morphological insensitivity in second language processing. *Applied Psycholinguistics, 25(4)*, 603-634.

Jiang, N. (2007). Selective integration of linguistic knowledge in adult second language learning. *Language learning, 57(1)*, 1-33.

Johnson, J. S., & Newport, E. L. (1991). Critical period effects on universal properties of language: The status of subjacency in the acquisition of a second language. *Cognition*, *39*(3), 215-258.

Kacetl, J., & Klímová, B. (2019). Use of smartphone applications in english language learning—A challenge for foreign language education. *Education Sciences, 9(3)*, 179.

Kaiser, J., & Lutzenberger, W. (2003). Induced gamma-band activity and human brain function. *The Neuroscientist, 9(6)*, 475-484.

Kaiser, J., Ripper, B., Birbaumer, N., & Lutzenberger, W. (2003). Dynamics of gamma-band activity in human magnetoencephalogram during auditory pattern working memory. *Neuroimage*, *20*(2), 816-827.

Kamiński, J., Brzezicka, A., Gola, M., & Wróbel, A. (2012). Beta band oscillations engagement in human alertness process. *International Journal of Psychophysiology, 85(1),* 125-128.

Kaan, E., Harris, A., Gibson, E., & Holcomb, P. (2000). The P600 as an index of syntactic integration difficulty. *Language and cognitive processes, 15*(2), 159-201.

Kandel, E. R., Koester, J. D., Mack, S. H., & Siegelbaum, S. (Eds.). (2021). *Principles of neural science*. New York: McGraw-hill.

Kanske, P., & Kotz, S. A. (2007). Concreteness in emotional words: ERP evidence from a hemifield study. *Brain research, 1148*, 138-148.

Karns, C. M., & Knight, R. T. (2009). Intermodal auditory, visual, and tactile attention modulates early stages of neural processing. *Journal of cognitive neuroscience, 21*(4), 669-683.

Kaschak, M. P., Jones, J. L., Carranza, J., & Fox, M. R. (2014). Embodiment and language comprehension. In *The Routledge handbook of embodied cognition* (pp. 136-144). Routledge.

Kelly, S. D., McDevitt, T., & Esch, M. (2009). Brief training with co-speech gesture lends a hand to word learning in a foreign language. *Language and cognitive processes*, *24(2),* 313-334.

Khanna, P., & Carmena, J. M. (2015). Neural oscillations: beta band activity across motor networks. *Current opinion in neurobiology, 32*, 60-67.

Khalsa, S. S., Rudrauf, D., Feinstein, J. S., & Tranel, D. (2009). The pathways of interoceptive awareness. *Nature neuroscience, 12*(12), 1494-1496.

Kiefer, M. (2005). Repetition-priming modulates category-related effects on event-related potentials: further evidence for multiple cortical semantic systems. *Journal of cognitive neuroscience*, *17*(2), 199-211.

Kiefer, M., & Pulvermüller, F. (2012). Conceptual representations in mind and brain: Theoretical developments, current evidence and future directions. *Cortex, 48*(7), 805-825.

Kiefer, M., Sim, E. J., Herrnberger, B., Grothe, J., & Hoenig, K. (2008). The sound of concepts: Four markers for a link between auditory and conceptual brain systems. *Journal of Neuroscience*, *28*(47), 12224-12230.

Kiefer, M., Trumpp, N., Herrnberger, B., Sim, E. J., Hoenig, K., & Pulvermüller, F. (2012). Dissociating the representation of action-and sound-related concepts in middle temporal cortex. *Brain and Language*, *122*(2), 120-125.

Kiehl, K. A., Liddle, P. F., Smith, A. M., Mendrek, A., Forster and, B. B., & Hare, R. D. (1999). Neural pathways involved in the processing of concrete and abstract words. *Human brain mapping, 7(4),* 225-233.

Klamer, S., Elshahabi, A., Lerche, H., Braun, C., Erb, M., Scheffler, K., & Focke, N. K. (2015). Differences between MEG and high-density EEG source localizations using a distributed source model in comparison to fMRI. *Brain topography, 28*(1), 87-94.

Klepp, A., Niccolai, V., Buccino, G., Schnitzler, A., & Biermann-Ruben, K. (2015). Language–motor interference reflected in MEG beta oscillations. *Neuroimage, 109,* 438-448.

Klimesch, W. (2012). Alpha-band oscillations, attention, and controlled access to stored information. *Trends in cognitive sciences, 16(12),* 606-617.

Kissler, J., Assadollahi, R., & Herbert, C. (2006). Emotional and semantic networks in visual word processing: insights from ERP studies. *Progress in brain research*, *156*, 147-183.

Kinoshita, S. (2006). Additive and interactive effects of word frequency and masked repetition in the lexical decision task. *Psychonomic Bulletin & Review, 13*(4), 668-673.

Kogan, B., Muñoz, E., Ibáñez, A., & García, A. M. (2020). Too late to be grounded? Motor resonance for action words acquired after middle childhood. *Brain and Cognition, 138,* 105509.

Kopf, J., Dresler, T., Reicherts, P., Herrmann, M. J., & Reif, A. (2013). The effect of emotional content on brain activation and the late positive potential in a word n-back task. *PloS one, 8*(9), e75598.

Kounios, J., & Holcomb, P. J. (1994). Concreteness effects in semantic processing: ERP evidence supporting dual-coding theory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20(4),* 804.

Kousta, S. T., Vigliocco, G., Vinson, D. P., Andrews, M., & Del Campo, E. (2011). The representation of abstract words: why emotion matters. *Journal of Experimental Psychology: General, 140(1),* 14.

Kraemer, D. J., Macrae, C. N., Green, A. E., & Kelley, W. M. (2005). Sound of silence activates auditory cortex. *Nature, 434(7030),* 158-158.

Krashen, S. D. (1976). Formal and informal linguistic environments in language acquisition and language learning. *Tesol Quarterly, 10(2),* 157-168.

Kroll, J. F. (1993). Accessing conceptual representations for words in a second language. *The bilingual lexicon, 53*, 481.

Kuhnke, P., Kiefer, M., & Hartwigsen, G. (2020). Task-dependent recruitment of modality-specific and multimodal regions during conceptual processing. *Cerebral Cortex, 30*(7), 3938-3959.

Kormos, J. (2019). How does vocabulary fit into theories of second language learning?. In *The Routledge handbook of vocabulary studies* (pp. 207-222). Routledge.

Kunzetsova, A., Brockhoff, P., & Christensen, R. (2017). lmerTest package: Tests in linear mixed effect models. . *Journal of Statistical Software*, *82*, 1-26.

Kutas, M., & Federmeier, K. D. (2000). Electrophysiology reveals semantic memory use in language comprehension. *Trends in cognitive sciences, 4*(12), 463-470.

Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: Finding meaning in the N400 component of the event related brain potential (ERP*). Annual review of psychology, 62*, 621.

Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science, 207*(4427), 203-205.

Lam, N. H., Schoffelen, J. M., Uddén, J., Hultén, A., & Hagoort, P. (2016). Neural activity during sentence processing as reflected in theta, alpha, beta, and gamma oscillations. *Neuroimage, 142*, 43-54.

Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological review, 104(2)*, 211.

Laufer, B. (2006). Comparing focus on form and focus on forms in second-language vocabulary learning. *Canadian Modern Language Review, 63*(1), 149-166.

Lee, C. Y., Liu, Y. N., & Tsai, J. L. (2012). The time course of contextual effects on visual word recognition. *Frontiers in Psychology, 3*, 285.

Lehmann, D., & Skrandies, W. (1984). Spatial analysis of evoked potentials in man--a review. *Progress in neurobiology, 23(3)*, 227-250.

Leicht, G., Björklund, J., Vauth, S., Mußmann, M., Haaf, M., Steinmann, S., ... & Mulert, C. (2021). Gamma-band synchronisation in a frontotemporal auditory information processing network. *Neuroimage*, *239*, 118307.

Lega, B. C., Jacobs, J., & Kahana, M. (2012). Human hippocampal theta oscillations and the formation of episodic memories. *Hippocampus, 22(4)*, 748-761.

Lewis, A. G., Schoffelen, J. M., Hoffmann, C., Bastiaansen, M., & Schriefers, H. (2017). Discourse-level semantic coherence influences beta oscillatory dynamics and the N400 during sentence comprehension. *Language, Cognition and Neuroscience*, *32(5)*, 601-617.

Li, P., Zhang, F., Yu, A., & Zhao, X. (2020). Language History Questionnaire (LHQ3): An enhanced tool for assessing multilingual experience. *Bilingualism: Language and Cognition*, *23(5)*, 938-944.

Light, G. A., Hsu, J. L., Hsieh, M. H., Meyer-Gomes, K., Sprock, J., Swerdlow, N. R., & Braff, D. L. (2006). Gamma band oscillations reveal neural network cortical coherence dysfunction in schizophrenia patients. *Biological psychiatry*, *60*(11), 1231-1240.

Long, M. H. (1990). Maturational constraints on language development. *Studies in second language acquisition*, *12*(3), 251-285.

Loonen, A. J., Schellekens, A. F., & Ivanova, S. A. (2016). Circuits regulating pleasure and happiness: a focus on addiction, beyond the ventral striatum. *Recent Advances in Drug Addiction Research and Clinical Applications, eds WM Meil and CL Ruby (Rijeka: InTech)*, 1-20.

Louwerse, M. M. (2008). Embodied relations are encoded in language. *Psychonomic Bulletin & Review, 15(4)*, 838-844.

Louwerse, M. M. (2018). Knowing the meaning of a word by the linguistic and perceptual company it keeps. *Topics in cognitive science, 10(3),* 573-589.

Luck, D., Danion, J. M., Marrer, C., Pham, B. T., Gounot, D., & Foucher, J. (2010). The right parahippocampal gyrus contributes to the formation and maintenance of bound information in working memory. *Brain and cognition, 72*(2), 255-263.

Luck, S. J., & Hillyard, S. A. (1994). Electrophysiological correlates of feature analysis during visual search. *Psychophysiology, 31*(3), 291-308.

Lupyan, G., & Winter, B. (2018). Language is more abstract than you think, or, why aren't languages more iconic?. *Philosophical Transactions of the Royal Society B: Biological Sciences, 373(1752),* 20170137.

Luu, P., Tucker, D. M., Englander, R., Lockfeld, A., Lutsep, H., & Oken, B. (2001). Localizing acute stroke-related eeg changes:: assessing the effects of spatial undersampling. *Journal of clinical Neurophysiology, 18*(4), 302-317.

Lynott, D., Connell, L., Brysbaert, M., Brand, J., & Carney, J. (2020). The Lancaster Sensorimotor Norms: multidimensional measures of perceptual and action strength for 40,000 English words. *Behavior Research Methods*, *52*(3), 1271-1291.

Lystad, R. P., & Pollard, H. (2009). Functional neuroimaging: a brief overview and feasibility for use in chiropractic research. *The Journal of the Canadian Chiropractic Association, 53*(1), 59.

Macedonia, M., & Klimesch, W. (2014). Long-term effects of gestures on memory for foreign language words trained in the classroom. *Mind, Brain, and Education, 8(2),* 74-88.

Macedonia, M., & Mueller, K. (2016). Exploring the neural representation of novel words learned through enactment in a word recognition task. *Frontiers in psychology, 7*, 953.

Macedonia, M., Müller, K., & Friederici, A. D. (2011). The impact of iconic gestures on foreign language word learning and its neural substrate. *Human brain mapping, 32(6),* 982-998.

Mahon, B. Z., & Caramazza, A. (2008). A critical look at the embodied cognition hypothesis and a new proposal for grounding conceptual content. *Journal of physiology-Paris, 102*(1-3), 59-70.

Makeig, S. (1993). Auditory event-related dynamics of the EEG spectrum and effects of exposure to tones. *Electroencephalography and clinical neurophysiology, 86(4),* 283-293.

Marieb, E. N., & Hoehn, K. (2007). *Human anatomy & physiology*. Pearson education.

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG-and MEG-data. *Journal of neuroscience methods*, *164*(1), 177-190.

Marschark, M., & Paivio, A. (1977). Integrative processing of concrete and abstract sentences. *Journal of verbal learning and verbal behavior, 16(2)*, 217-231.

Matheson, H. E., & Barsalou, L. W. (2018). Embodiment and grounding in cognitive neuroscience. *Stevens' handbook of experimental psychology and cognitive neuroscience, 3*, 1-27.

Mayer, K. M., Yildiz, I. B., Macedonia, M., & von Kriegstein, K. (2015). Visual and motor cortices differentially support the translation of foreign language words. *Current Biology, 25(4),* 530-535.

Mazzuca, C., Lugli, L., Benassi, M., Nicoletti, R., & Borghi, A. M. (2018). Abstract, emotional and concrete concepts and the activation of mouth-hand effectors. PeerJ, 6, e5987.

McCandliss, B. D., Cohen, L., & Dehaene, S. (2003). The visual word form area: expertise for reading in the fusiform gyrus. *Trends in cognitive sciences, 7*(7), 293-299.

Mégevand, P., Spinelli, L., Genetti, M., Brodbeck, V., Momjian, S., Schaller, K., ... & Seeck, M. (2014). Electric source imaging of interictal activity accurately localises the seizure onset zone. *Journal of Neurology, Neurosurgery & Psychiatry, 85*(1), 38-43.

Meinhardt, J., & Pekrun, R. (2003). Attentional resource allocation to emotional events: An ERP study. *Cognition and Emotion, 17*(3), 477-500.

Mellem, M. S., Friedman, R. B., & Medvedev, A. V. (2013). Gamma-and theta-band synchronization during semantic priming reflect local and long-range lexical–semantic networks. *Brain and language, 127(3),* 440-451.

Melynyte, S., Wang, G. Y., & Griskova-Bulanova, I. (2018). Gender effects on auditory P300: A systematic review. *International Journal of Psychophysiology, 133*, 55-65.

Meyer, P., Mecklinger, A., Grunwald, T., Fell, J., Elger, C. E., & Friederici, A. D. (2005). Language processing within the human medial temporal lobe. *Hippocampus, 15*(4), 451-459.

Miceli, A., Wauthia, E., Lefebvre, L., Ris, L., & Simoes Loureiro, I. (2021). Perceptual and Interoceptive Strength Norms for 270 French Words. *Frontiers in Psychology*, *12*, 2018.

Michel, C. M. (2019). High-resolution EEG. *Handbook of Clinical Neurology, 160*, 185-201.

Michel, C. M., & He, B. (2019). EEG source localization. *Handbook of clinical neurology, 160*, 85-101.

Miklashevsky, A. (2018). Perceptual experience norms for 506 Russian nouns: Modality rating, spatial localization, manipulability, imageability and other variables. *Journal of psycholinguistic research, 47*(3), 641-661.

Miltner, W. H., Braun, C., Arnold, M., Witte, H., & Taub, E. (1999). Coherence of gamma-band EEG activity as a basis for associative learning. *Nature, 397(6718)*, 434-436.

Misra, M., & Holcomb, P. J. (2003). Event–related potential indices of masked repetition priming. *Psychophysiology, 40*(1), 115-130.

Mkrtychian, N., Blagovechtchenski, E., Kurmakaeva, D., Gnedykh, D., Kostromina, S., & Shtyrov, Y. (2019). Concrete vs. abstract semantics: from mental representations to functional brain mapping. *Frontiers in human neuroscience, 267*.

Monaco, E., Jost, L. B., Gygax, P. M., & Annoni, J. M. (2019). Embodied semantics in a second language: Critical review and clinical implications. *Frontiers in Human Neuroscience, 13*, 110.

Moreno, I., De Vega, M., León, I., Bastiaansen, M., Lewis, A. G., & Magyari, L. (2015). Brain dynamics in the comprehension of action-related language. A time-frequency analysis of mu rhythms. *Neuroimage, 109*, 50-62.

Mori, K., Manabe, H., Narikiyo, K., & Onisawa, N. (2013). Olfactory consciousness and gamma oscillation couplings across the olfactory bulb, olfactory cortex, and orbitofrontal cortex. *Frontiers in psychology, 4*, 743.

Morillon, B., Lehongre, K., Frackowiak, R. S., Ducorps, A., Kleinschmidt, A., Poeppel, D., & Giraud, A. L. (2010). Neurophysiological origin of human brain asymmetry for speech and language. *Proceedings of the National Academy of Sciences, 107*(43), 18688-18693.

Moseley, R., Carota, F., Hauk, O., Mohr, B., & Pulvermüller, F. (2012). A role for the motor system in binding abstract emotional meaning. *Cerebral cortex, 22*(7), 1634-1647.

Muraki, E. J., Cortese, F., Protzner, A. B., & Pexman, P. M. (2020). Heterogeneity in abstract verbs: An ERP study. *Brain and language, 211*, 104863.

Müller, N. G., Strumpf, H., Scholz, M., Baier, B., & Melloni, L. (2013). Repetition suppression versus enhancement—it's quantity that matters. *Cerebral cortex, 23*(2), 315-322.

Myachykov, A., & Fischer, M. H. (2019). A hierarchical view of abstractness: Grounded, embodied, and situated aspects. Comment on" Words as social tools: Language, sociality and inner grounding in abstract concepts" by Anna M. Borghi et al. *Physics of Life Reviews, 29*, 161-163.

Nakayama, M., & Lupker, S. J. (2018). Is there lexical competition in the recognition of L2 words for different-script bilinguals? An examination using masked priming with Japanese-English bilinguals. *Journal of Experimental Psychology: Human Perception and Performance, 44(8)*, 1168.

Nation, I. (2013a). Learning words from context. In *Learning Vocabulary in Another Language* (Cambridge Applied Linguistics, pp. 348-388). Cambridge: Cambridge University Press.

Nation, I. (2013b). The goals of vocabulary learning. In *Learning Vocabulary in Another Language* (Cambridge Applied Linguistics, pp. 9-43).

Cambridge: Cambridge University Press.

National Institute of the Korean Language. (n.d.). 디스크. In *NIKL Korean Dictionary*. Retrieved December 22, 2022. from https://stdict.korean.go.kr/

Newton, J. (2019). Approaches to learning vocabulary inside the classroom. In The *Routledge handbook of vocabulary studies* (pp. 255-270). Routledge.

Nieuwland, M. S. (2019). Do 'early' brain responses reveal word form prediction during language comprehension? A critical review. *Neuroscience & Biobehavioral Reviews, 96*, 367-400.

Noppeney, U., & Price, C. J. (2004). Retrieval of abstract semantics. *Neuroimage, 22(1)*, 164-170.

O'Grady, W. (2008). Innateness, universal grammar, and emergentism. *Lingua*, *118*(4), 620-631.

Öttl, B., Dudschig, C., & Kaup, B. (2017). Forming associations between language and sensorimotor traces during novel word learning. *Language and Cognition*, *9*(1), 156-171.

Pascual-Marqui, R. D. (2002). Standardized low-resolution brain electromagnetic tomography (sLORETA): technical details. *Methods Find Exp Clin Pharmacol, 24*(Suppl D), 5-12.

Pascual-Marqui, R. D., Esslen, M., Kochi, K., & Lehmann, D. (2002). Functional imaging with low-resolution brain electromagnetic tomography (LORETA): a review. *Methods and findings in experimental and clinical pharmacology, 24*(Suppl C), 91-95.

Pascual-Marqui, R. D., Lehmann, D., Koukkou, M., Kochi, K., Anderer, P., Saletu, B., ... & Kinoshita, T. (2011). Assessing interactions in the brain with exact low-resolution electromagnetic tomography. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *369*(1952), 3768-3784.

Paivio, A. (1986). *Mental representations: A dual coding approach.* New York: Oxford University Press.

Palazova, M., Mantwill, K., Sommer, W., & Schacht, A. (2011). Are effects of emotion in single words non-lexical? Evidence from event-related brain potentials. *Neuropsychologia*, *49*(9), 2766-2775.

Pantev, C. (1995). Evoked and induced gamma-band activity of the human cortex. *Brain topography*, *7*(4), 321-330.

Pantev, C., Elbert, T., Makeig, S., Hampson, S., Eulitz, C., & Hoke, M. (1993). Relationship of transient and steady-state auditory evoked fields. *Electroencephalography and Clinical Neurophysiology/Evoked Potentials Section*, *88*(5), 389-396.

Patterson, K., Nestor, P. J., & Rogers, T. T. (2007). Where do you know what you know? The representation of semantic knowledge in the human brain. *Nature reviews neuroscience*, *8*(12), 976-987.

Pavlenko, A. (2009). Conceptual representation in the bilingual lexicon and second language vocabulary learning. *The bilingual mental lexicon: Interdisciplinary approaches, 70*.

Pavlenko, A. (2012). Affective processing in bilingual speakers: Disembodied cognition?. *International Journal of Psychology*, *47*(6), 405-428.

Penolazzi, B., Hauk, O., & Pulvermüller, F. (2007). Early semantic context integration and lexical access as revealed by event-related brain potentials. *Biological psychology*, *74*(3), 374-388.

Peña, M., & Melloni, L. (2012). Brain oscillations during spoken sentence processing. *Journal of cognitive neuroscience, 24*(5), 1149-1164.

Perani, D., Paulesu, E., Galles, N. S., Dupoux, E., Dehaene, S., Bettinardi, V., ... & Mehler, J. (1998). The bilingual brain. Proficiency and age of acquisition of the second language. *Brain: a journal of neurology, 121*(10), 1841-1852.

Pexman, P. M., Hargreaves, I. S., Edwards, J. D., Henry, L. C., & Goodyear, B. G. (2007). Neural correlates of concreteness in semantic categorization. *Journal of Cognitive Neuroscience, 19(8),* 1407-1419.

Pfurtscheller, G., & Da Silva, F. L. (1999). Event-related EEG/MEG synchronization and desynchronization: basic principles. *Clinical neurophysiology, 110(11),* 1842-1857.

Phillips, M. L., Drevets, W. C., Rauch, S. L., & Lane, R. (2003). Neurobiology of emotion perception I: The neural basis of normal emotion perception. *Biological psychiatry, 54*(5), 504-514.

Picton, T. W. (1992). The P300 wave of the human event-related potential. *Journal of clinical neurophysiology, 9(4),* 456-479.

Poeppel, D. (2014). The neuroanatomic and neurophysiological infrastructure for speech and language. *Current opinión in neurobiology, 28*, 142-149.

Poeppel, D., & Hickok, G. (2004). Towards a new functional anatomy of language. *Cognition, 92*(1-2), 1-12.

Pogosyan, A., Gaynor, L. D., Eusebio, A., & Brown, P. (2009). Boosting cortical activity at beta-band frequencies slows movement in humans. *Current biology, 19(19),* 1637-1641.

Polich, J. (2007). Updating P300: an integrative theory of P3a and P3b. *Clinical* *neurophysiology, 118*(10), 2128-2148.

Polich, J., & Donchin, E. (1988). P300 and the word frequency effect. *Electroencephalography and clinical neurophysiology, 70*(1), 33-45.

Ponari, M., Norbury, C. F., & Vigliocco, G. (2018). Acquisition of abstract concepts is influenced by emotional valence. *Developmental science, 21(2)*, e12549.

Posner, M. I., & Abdullaev, Y. G. (1999). Neuroanatomy, circuitry and plasticity of word reading. *NeuroReport: For Rapid Communication of Neuroscience Research, 10*(3), R12–R23.

Popp, M., Trumpp, N. M., & Kiefer, M. (2016). Feature-specific event-related potential effects to action-and sound-related verbs during visual word recognition. *Frontiers in Human Neuroscience*, *10*, 637.

Preissl, H., Pulvermüller, F., Lutzenberger, W., & Birbaumer, N. (1995). Evoked potentials distinguish between nouns and verbs. *Neuroscience Letters, 197*(1), 81-83.

Prystauka, Y., & Lewis, A. G. (2019). The power of neural oscillations to inform sentence comprehension: A linguistic perspective. *Language and Linguistics Compass*, *13*(9), e12347.

Pulvermüller, F. (1996). Hebb's concept of cell assemblies and the psychophysiology of word processing. *Psychophysiology, 33(4)*, 317-333.

Pulvermüller, F. (1999). Words in the brain's language. *Behavioral and brain sciences*, *22(2)*, 253-279.

Pulvermüller, F. (2005). Brain mechanisms linking language and action. *Nature reviews neuroscience*, *6(7)*, 576-582.

Pulvermüller, F., Eulitz, C., Pantev, C., Mohr, B., Feige, B., Lutzenberger, W., ... & Birbaumer, N. (1996). High-frequency cortical responses reflect lexical processing: an MEG study. *Electroencephalography and clinical Neurophysiology, 98*(1), 76-85.

Pulvermüller, F., Hauk, O., Nikulin, V. V., & Ilmoniemi, R. J. (2005). Functional links between motor and language systems. *European Journal of Neuroscience, 21*(3), 793-797.

Pulvermüller, F., Kherif, F., Hauk, O., Mohr, B., & Nimmo-Smith, I. (2009). Distributed cell assemblies for general lexical and category-specific semantic processing as revealed by fMRI cluster analysis. *Human brain mapping, 30(12),* 3837-3850.

Pulvermüller, F., Lutzenberger, W., & Preissl, H. (1999). Nouns and verbs in the intact brain: Evidence from event-related potentials and high-frequency cortical responses. *Cerebral cortex*, *9*(5), 497-506.

Qian, D. (1999). Assessing the roles of depth and breadth of vocabulary knowledge in reading comprehension. *Canadian modern language review, 56*(2), 282-308.

Qian, D. D., & Lin, L. H. (2019). The relationship between vocabulary knowledge and language proficiency. In The *Routledge handbook of vocabulary studies* (pp. 66-80). Routledge.

Raju, H., & Tadi, P. (2020). Neuroanatomy, Somatosensory Cortex. *StatPearls [Internet]*.

Raslau, F. D., Mark, I. T., Klein, A. P., Ulmer, J. L., Mathews, V., & Mark, L. P. (2015). Memory part 2: the role of the medial temporal lobe. *American Journal of Neuroradiology, 36*(5), 846-849.

Rasmussen, T., & Milner, B. (1977). The role of early left-brain injury in determining lateralization of cerebral speech functions. *Annals of the New York Academy of Sciences*, *299*, 355-369.

R Core Team (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Reilly, J., Flurie, M., & Peelle, J. E. (2020). The English lexicon mirrors functional brain activation for a sensory hierarchy dominated by vision and audition: Point-counterpoint. *Journal of neurolinguistics, 55*, 100895.

Ribary, U. (2005). Dynamics of thalamo-cortical network oscillations and human perception. *Progress in brain research, 150*, 127-142.

Rizzolatti, G., & Craighero, L. (2004). *The mirror-neuron system.* Annu. Rev. Neurosci., 27, 169-192.

Rodriguez, E., George, N., Lachaux, J. P., Martinerie, J., Renault, B., & Varela, F. J. (1999). Perception's shadow: long-distance synchronization of human brain activity. *Nature, 397*(6718), 430-433.

Rolls, E. T. (2004). The functions of the orbitofrontal cortex. *Brain and cognition, 55(1)*, 11-29.

Röschke, J., & Wagner, P. (2003). A confirmatory study on the mechanisms behind reduced P300 waves in depression. *Neuropsychopharmacology, 28*(1), S9-S12.

Rugg, M. D., & Curran, T. (2007). Event-related potentials and recognition memory. *Trends in cognitive sciences, 11*(6), 251-257.

Rugg, M. D., Mark, R. E., Walla, P., Schloerscheidt, A. M., Birch, C. S., & Allan, K. (1998). Dissociation of the neural correlates of implicit and explicit memory. *Nature*, *392*(6676), 595-598.

Rugg, M. D., & Nieto-Vegas, M. (1999). Modality-specific effects of immediate word repetition: Electrophysiological evidence. *NeuroReport, 10*(12), 2661-2664.

Rushworth, M. F., Behrens, T. E. J., Rudebeck, P. H., & Walton, M. E. (2007). Contrasting roles for cingulate and orbitofrontal cortex in decisions and social behaviour. *Trends in cognitive sciences, 11(4)*, 168-176.

Sakreida, K., Scorolli, C., Menz, M. M., Heim, S., Borghi, A. M., & Binkofski, F. (2013). Are abstract action words embodied? An fMRI investigation at the interface between language and motor cognition. *Frontiers in human neuroscience, 7*, 125.

Schacht, A., & Sommer, W. (2009). Time course and task dependence of emotion effects in word processing. *Cognitive, Affective, & Behavioral Neuroscience, 9*(1), 28-43.

Schaller, F., Weiss, S., & Müller, H. M. (2017). EEG beta-power changes reflect motor involvement in abstract action language processing. *Brain and language, 168*, 95-105.

Schupp, H. T., Flaisch, T., Stockburger, J., & Junghöfer, M. (2006). Emotion and attention: event-related brain potential studies. *Progress in brain research, 156*, 31-51.

Schwanenflugel, P. J., & Stowe, R. W. (1989). Context availability and the processing of abstract and concrete words in sentences. *Reading Research Quarterly*, 114-126.

Schwanenflugel, P. J., Harnishfeger, K. K., & Stowe, R. W. (1988). Context availability and lexical decisions for abstract and concrete words.

Journal of Memory and Language, *27(5)*, 499-520.

Schweinberger, S. R., Pfütze, E. M., & Sommer, W. (1995). Repetition priming and associative priming of face recognition: Evidence from event-related potentials. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21(3),* 722.

Scott, G. G., O'Donnell, P. J., Leuthold, H., & Sereno, S. C. (2009). Early emotion word processing: Evidence from event-related potentials. *Biological psychology, 80(1),* 95-104.

Searle, J. R. (1980). Minds, brains, and programs. *Behavioral and brain sciences*, *3*(3), 417-424.

Seabold, S., & Perktold, J. (2010, June). Statsmodels: Econometric and statistical modeling with python. In *Proceedings of the 9th Python in Science Conference* (Vol. 57, No. 61, pp. 10-25080).

Shriram, R., Sundhararajan, M., & Daimiwal, N. (2013). EEG based cognitive workload assessment for maximum efficiency. *IOSR Journal of Electronics and Communication Engineering*, *7*, 34-38.

Singer, W., & Gray, C. M. (1995). Visual feature integration and the temporal correlation hypothesis. *Annual review of neuroscience, 18(1)*, 555-586.

Smith, E. E., Shoben, E. J., & Rips, L. J. (1974). Structure and process in semantic memory: A featural model for semantic decisions. *Psychological review, 81*(3), 214.

Speed, L. J., & Brybaert, M. (2021). Dutch sensory modality norms. *Behavior Research Methods*, 1-13.

Speed, L. J., & Majid, A. (2017). Dutch modality exclusivity norms: Simulating perceptual modality in space. *Behavior research methods*, *49*(6), 2204-

2218.

Speer, R., Chin, J., Lin, A., Jewett, S., & Nathan, L., (2018). LuminosoInsight/wordfreq: v2.2 (v2.2). Zenodo. https://doi.org/10.5281/zenodo.1443582

Squire, L. R., Stark, C. E., & Clark, R. E. (2004). The medial temporal lobe. *Annu. Rev. Neurosci*., 27, 279-306.

Steinmann, S., Leicht, G., Ertl, M., Andreou, C., Polomac, N., Westerhausen, R., ... & Mulert, C. (2014). Conscious auditory perception related to long-range synchrony of gamma oscillations. *Neuroimage, 100*, 435-443.

Strüber, D., Basar-Eroglu, C., Hoff, E., & Stadler, M. (2000). Reversal-rate dependent differences in the EEG gamma-band during multistable visual perception. *International journal of psychophysiology, 38*(3), 243-252.

Sueur, J. (2018). Introduction to Frequency Analysis: The Fourier Transformation. In *Sound Analysis and Synthesis with R* (pp. 213-245). Springer, Cham.

Tainturier, M. J., Tamminen, J., & Thierry, G. (2005). Age of acquisition modulates the amplitude of the P300 component in spoken word recognition. *Neuroscience letters, 379*(1), 17-22.

Tallon-Baudry, C., Bertrand, O., Delpuech, C., & Pernier, J. (1997). Oscillatory γ-band (30–70 Hz) activity induced by a visual search task in humans. *Journal of Neuroscience, 17*(2), 722-734.

Tallon-Baudry, C., Bertrand, O., Peronnet, F., & Pernier, J. (1998). Induced γ-band activity during the delay of a visual short-term memory task in humans. *Journal of Neuroscience*, *18*(11), 4244-4254.

Tanji, K., Suzuki, K., Delorme, A., Shamoto, H., & Nakasato, N. (2005). High-frequency γ-band activity in the basal temporal cortex during picture-naming and lexical-decision tasks. *Journal of Neuroscience, 25*(13), 3287-3293.

Tesche, C. D., & Karhu, J. (2000). Theta oscillations index human hippocampal activation during a working memory task. *Proceedings of the National Academy of Sciences, 97(2)*, 919-924.

Tettamanti, M., Buccino, G., Saccuman, M. C., Gallese, V., Danna, M., Scifo, P., ... & Perani, D. (2005). Listening to action-related sentences activates fronto-parietal motor circuits. *Journal of cognitive neuroscience, 17(2)*, 273-281.

Tiitinen, H. T., Sinkkonen, J., Reinikainen, K., Alho, K., Lavikainen, J., & Näätänen, R. (1993). Selective attention enhances the auditory 40-Hz transient response in humans. *Nature*, *364*(6432), 59-60.

Tremblay, P., & Dick, A. S. (2016). Broca and Wernicke are dead, or moving past the classic model of language neurobiology. *Brain and language*, *162*, 60-71.

Trumpp, N. M., Kliese, D., Hoenig, K., Haarmeier, T., & Kiefer, M. (2013). Losing the sound of concepts: Damage to auditory association cortex impairs the processing of sound-related concepts. Cortex, 49(2), 474-486.

Trumpp, N. M., Traub, F., & Kiefer, M. (2013). Masked priming of conceptual features reveals differential brain activation during unconscious access to conceptual action and sound information. PLoS One, 8(5), e65910.

Trumpp, N. M., Traub, F., Pulvermüller, F., & Kiefer, M. (2014). Unconscious automatic brain activation of acoustic and action-related conceptual features during masked repetition priming. Journal of cognitive neuroscience, 26(2), 352-364.

Tsao, A., Sugar, J., Lu, L., Wang, C., Knierim, J. J., Moser, M. B., & Moser, E. I. (2018). Integrating time from experience in the lateral entorhinal cortex. *Nature*, *561*(7721), 57-62.

Tso, I. F., Rutherford, S., Fang, Y., Angstadt, M., & Taylor, S. F. (2018). The "social brain" is highly sensitive to the mere presence of social information: An automated meta-analysis and an independent study. *PloS one, 13(5),* e0196503.

Turk-Browne, N. B., Yi, D. J., Leber, A. B., & Chun, M. M. (2007). Visual quality determines the direction of neural repetition effects. *Cerebral Cortex, 17*(2), 425-433.

Ullman, M. T. (2001a). A neurocognitive perspective on language: The declarative/procedural model. *Nature reviews neuroscience, 2(10*), 717-726.

Ullman, M. T. (2001b). The declarative/procedural model of lexicon and grammar. *Journal of psycholinguistic research, 30(1)*, 37-69.

Ullman, M. T. (2004). Contributions of memory circuits to language: The declarative/procedural model. *Cognition, 92(1-2)*, 231-270.

Ullman, M. T. (2015). The declarative/procedural model: A neurobiologically motivated theory of first and second language. In *Theories in second language acquisition* (pp. 147-172). Routledge.

Vaidya, A. R., Pujara, M. S., Petrides, M., Murray, E. A., & Fellows, L. K. (2019). Lesion studies in contemporary neuroscience. *Trends in cognitive sciences*, *23*(8), 653-671.

Vallat, R. (2018). Pingouin: statistics in Python. *Journal of Open Source Software, 3(31),* 1026, https://doi.org/10.21105/joss.01026

Van den Bussche, E., Van den Noortgate, W., & Reynvoet, B. (2009). Mechanisms of masked priming: a meta-analysis. *Psychological bulletin*, *135*(3), 452.

Vanhoutte, S., Strobbe, G., van Mierlo, P., Cosyns, M., Batens, K., Corthals, P., ... & Santens, P. (2015). Early lexico-semantic modulation of motor related areas during action and non-action verb processing. *Journal of neurolinguistics, 34*, 65-82.

van Hell, J.G., & De Groot, A. M. (1998). Disentangling context availability and concreteness in lexical decision and word translation. *The Quarterly Journal of Experimental Psychology: Section A, 51(1)*, 41-63.

Van Heuven, W. J., Mandera, P., Keuleers, E., & Brysbaert, M. (2014). SUBTLEX-UK: A new and improved word frequency database for British English. *Quarterly journal of experimental psychology, 67*(6), 1176-1190.

VanPatten, B., & Williams, J. (2015a). Introduction: The nature of theories. In *Theories in second language acquisition* (pp. 13-28). Routledge.

VanPatten, B., & Williams, J. (2015b). Early Theories in SLA. In *Theories in Second Language Acquisition* (pp. 29-45). Routledge.

Vasanta, D., Raju, S. B., Suvarna, A., Patel, J., Naidu, Y. V., Sireesha, J., & Nigam, R. (2011). Action verbs and body parts: a cross-linguistic study. *International Journal of Mind Brain and Cognition*, 2(1-2), 19-38.

Vergallito, A., Petilli, M. A., & Marelli, M. (2020). Perceptual modality norms for 1,121 Italian words: A comparison with concreteness and imageability scores and an analysis of their impact in word processing tasks. *Behavior research methods*, 1-18.

Vermeulen, N., Niedenthal, P. M., & Luminet, O. (2007). Switching between sensory and affective systems incurs processing costs. *Cognitive*

Science, 31(1), 183-192.

Vidal, K. (2011). A comparison of the effects of reading and listening on incidental vocabulary acquisition. *Language Learning, 61*(1), 219-258.

Vigliocco, G., Kousta, S. T., Della Rosa, P. A., Vinson, D. P., Tettamanti, M., Devlin, J. T., & Cappa, S. F. (2014). The neural representation of abstract words: the role of emotion. *Cerebral Cortex*, *24*(7), 1767-1777.

Villani, C., D'Ascenzo, S., Borghi, A. M., Roversi, C., Benassi, M., & Lugli, L. (2021). Is justice grounded? How expertise shapes conceptual representation of institutional concepts. *Psychological Research*, 1-17.

Villani, C., Lugli, L., Liuzza, M. T., & Borghi, A. M. (2019). Varieties of abstract concepts and their multiple dimensions. *Language and Cognition, 11(3),* 403-430.

Villani, C., Lugli, L., Liuzza, M. T., Nicoletti, R., & Borghi, A. M. (2021). Sensorimotor and interoceptive dimensions in concrete and abstract concepts. *Journal of memory and language, 116*, 104173.

Villardita, C., Grioli, S., & Quattropani, M. C. (1988). Concreteness/abstractness of stimulus-words and semantic clustering in right brain-damaged patients. *Cortex*, *24(4),* 563-571.

Visser, M., Jefferies, E., & Lambon Ralph, M. A. (2010). Semantic processing in the anterior temporal lobes: a meta-analysis of the functional neuroimaging literature. *Journal of cognitive neuroscience, 22*(6), 1083-1094.

Vukovic, N. (2013). When words get physical: evidence for proficiency-modulated somatotopic motor interference during second language comprehension. In *Proceedings of the Annual Meeting of the Cognitive Science Society*

(Vol. 35, No. 35).

Vukovic, N., & Shtyrov, Y. (2014). Cortical motor systems are involved in second-language comprehension: evidence from rapid mu-rhythm desynchronisation. *NeuroImage*, *102*, 695-703.

Vukovic, N., & Williams, J. N. (2014). Automatic perceptual simulation of first language meanings during second language sentence processing in bilinguals. *Acta psychologica*, *145*, 98-103.

Wang, J., Conder, J. A., Blitzer, D. N., & Shinkareva, S. V. (2010). Neural representation of abstract and concrete concepts: A meta-analysis of neuroimaging studies. *Human brain mapping, 31*(10), 1459-1468.

Wang, X., Shangguan, C., & Lu, J. (2019). Time course of emotion effects during emotion-label and emotion-laden word processing. *Neuroscience letters*, *699*, 1-7.

Warrington, E. K., & McCarthy, R. A. (1987). Categories of knowledge: Further fractionations and an attempted integration. *Brain*, *110*(5), 1273-1296.

Webb, S. (2007). The effects of repetition on vocabulary knowledge. *Applied linguistics*, *28*(1), 46-65.

Webb, S. (2019a). Introduction. In The *Routledge handbook of vocabulary studies* (pp. 1-12). Routledge.

Webb, S. (2019b). *Incidental vocabulary learning*. In The Routledge handbook of vocabulary studies (pp. 225-239). Routledge.

Weiss, S., & Mueller, H. M. (2012). “Too many betas do not spoil the broth”: the role of beta brain oscillations in language processing. *Frontiers in psychology, 3*, 201.

Welcome, S. E., Paivio, A., McRae, K., & Joanisse, M. F. (2011). An electrophysiological study of task demands on concreteness effects: evidence for dual coding theory. *Experimental brain research, 212*(3), 347-358.

West, W. C., & Holcomb, P. J. (2000). Imaginal, semantic, and surface-level processing of concrete and abstract words: an electrophysiological investigation. *Journal of Cognitive Neuroscience, 12*(6), 1024-1037.

Wianda, E., & Ross, B. (2019). The roles of alpha oscillation in working memory retention. *Brain and behavior, 9(4),* e01263.

Wilson-Mendenhall, C. D., Henriques, A., Barsalou, L. W., & Barrett, L. F. (2019). Primary interoceptive cortex activity during simulated experiences of the body. *Journal of cognitive neuroscience, 31*(2), 221-235.

Wilson-Mendenhall, C. D., Simmons, W. K., Martin, A., & Barsalou, L. W. (2013). Contextual processing of abstract concepts reveals neural representations of nonlinguistic semantic content. *Journal of cognitive neuroscience, 25(6),* 920-935.

Yanagisawa, A., & Webb, S. (2019). Measuring depth of vocabulary knowledge. In *The Routledge handbook of vocabulary studies* (pp. 371-386). Routledge.

Yang, H., Laforge, G., Stojanoski, B., Nichols, E. S., McRae, K., & Köhler, S. (2019). Late positive complex in event-related potentials tracks memory signals when they are decision relevant. *Scientific reports, 9*(1), 1-15.

Xu, M., Li, D., & Li, P. (2021). Brain decoding in multiple languages: Can cross-language brain decoding work?. *Brain and Language*, *215,* 104922.

Xue, J., Marmolejo-Ramos, F., & Pei, X. (2015). The linguistic context effects on the processing of body–object interaction words: An ERP study on second language learners. *Brain research*, *1613*, 37-48.

Zhang, Q., Guo, C. Y., Ding, J. H., & Wang, Z. Y. (2006). Concreteness effects in the processing of Chinese words. *Brain and Language*, *96*(1), 59-68.

Zhang, Z. G., Hu, L., Hung, Y. S., Mouraux, A., & Iannetti, G. D. (2012). Gamma-band oscillations in the primary somatosensory cortex—a direct and obligatory correlate of subjective pain intensity. *Journal of Neuroscience*, *32*(22), 7429-7438.

Zang, B., Lin, Y., Liu, Z., & Gao, X. (2021). A deep learning method for single-trial EEG classification in RSVP task based on spatiotemporal features of ERPs. *Journal of Neural Engineering*, *18*(4), 0460c8.

Zhang, J., Wu, C., Yuan, Z., & Meng, Y. (2020). Different early and late processing of emotion-label words and emotion-laden words in a second language: An ERP study. *Second Language Research*, *36*(3), 399-412.

Zhao, T., Huang, Y., Chen, D., Jiao, L., Marmolejo-Ramos, F., Wang, R., & Xie, J. (2020). The modality switching costs of Chinese–English bilinguals in the processing of L1 and L2. *Quarterly Journal of Experimental Psychology*, *73*(3), 396-412.

Zehr, J., & Schwarz, F. (2018). PennController for Internet Based Experiments (IBEX). https://doi.org/10.17605/OSF.IO/MD832

Zwaan, R. A., & Madden, C. J. (2005). Embodied sentence comprehension. *Grounding cognition: The role of perception and action in memory, language, and thinking*, 224-245.

국문초록

제2언어처리에서의 추상 개념의 체화된 인지: 단어 감각 지수와 감각 특정 신경반응에 대한 EEG 분석

본 논문은 제2언어처리에서 추상개념의 체화된 인지를 단어 감각 지수와 감각 특정 신경반응에 대한 뇌파측정(EEG)실험을 통해 탐구하고자 한다. 체화된 인지 이론에 따르면 언어의 이해는 오직 언어특수영역에 의해서 이루어지는 것이 아니라 지각, 운동, 그리고 감정을 처리하는 동일한 뇌의 영역의 활성화에 의해서 이루어진다. 이런 가정은 언어처리 동안에 관련된 감각-운동영역이 활성화됨을 보여주는 연구들에 의해서 증명되어 왔다. 그러나 이런 연구들은 주로 운동관련 단어에 국한되어 왔고 제2언어사용자와 같은 다른 종류의 언어사용자에게는 수행되지 않았다. 본 논문은 이런 한계점을 다루고자 아직 체화된 인지 이론의 틀에서 완전히 탐구되지 않은 추상개념에 초점을 두어 제2언어 학습자의 언어처리에서의 체화된 인지를 탐구하고자 하였다.

본 논문은 먼저 추상개념이 구체개념과 마찬가지로 감각-운동 영역에 체화되었는지를 조사한다. 추상개념은 물리적인 대상물을 가지고 있지 않기 때문에 감각운동 경험을 통해서 이해되지 않는 것으로 여겨져 왔다. 몇몇의 연구는 최근에 감각-운동 영역의 활성화를 통해 이런

추상개념을 설명하기 위해서 내부 신체의 감각 지각을 의미하는 내부수용감각을 도입하는 것을 시도하였다. 본 연구는 실험1에서 내부수용감각을 포함한 단어 감각 지수를 평가하고 실험2에서 청각관련 단어와 내부수용감각 단어 간의 감각 특정 신경반응을 비교하여 이러한 문제를 연구하고자 한다.

본 논문은 또한 제2언어처리가 제1언어처리와 마찬가지로 감각운동영역에 체화되는지를 탐구하고자 한다. 본 연구의 실험참여자들은 외국어 환경에서 영어를 배운 제2언어 학습자들이다. 이런 환경은 언어 학습과 사용의 측면에서 제1언어사용자의 환경에 비해 질적이고 양적인 차이가 있는데, 이런 환경에서 제2언어학습자는 환경과의 직접적인 접촉이 적으며 종종 그들의 모국어에 의한 중재를 통해 언어를 학습하게 된다. 언어의 체화는 언어와 환경간의 반복적인 연결을 통해 형성된다고 가정되기 때문에, 이러한 환경에서의 차이는 제2언어처리의 체화에 결정적인 영향을 미칠지도 모른다. 이를 탐구하기 위해 본 연구는 실험1에서 제2언어사용자와 모국어사용자의 단어감각지수의 분포를 비교하고 실험2에서 감각 특정 신경 반응에서의 차이를 탐구한다.

이런 연구 주제를 탐구하기 위해서 두 개의 실험이 수행되었다. 실험1의 목적은 한국어, 영어, L2영어의 단어 감각 지수를 구축하고 그것들을 서로 비교하는 것이었다. 단어 감각 지수는 단어의 감각 프로필로, 어떻게 단어가 오감과 내부수용감각을 통해서 경험되는지를 측정한다. 이것은 후속 연구들의 자료로 사용될 수 있을 뿐만 아니라

제1언어와 제2언어사용자들이 언어를 내부수용감각을 포함한 인간 지각 경험에 얼마나 잘 연결지을 수 있는지를 탐구할 수 있게 해준다. 실험2는 청각단어와 내부수용감각 단어가 각기 다른 신경반응을 유도하는지 그리고 이러한 반응이 제1언어사용자와 제2언어사용자 간에 어떻게 다른지를 연구하는 것을 목표로 한다. 이를 위해 뇌파측정 실험이 수행되었는데, 이 실험에서 한국어 모국어 화자는 청각 단어와 내부수용감각 단어를 포함하는 어휘판단과제를 한국어와 영어로 수행하였다. 체화된 인지 이론에 따르면, 청각단어는 청각관련 영역에서 처리되고 내부수용감각은 내부수용감각관련 영역에서 처리되기 때문에 두 감각의 단어 사이에서 구별되는 신경반응이 관찰될 것이다.

실험1과 실험2의 결과는 추상개념도 또한 감각-운동 영역에 체화되었음을 드러내었다. 실험1은 내부수용감각과 추상성 사이에서 유의미한 상관관계가 있음을 발견하였다. 이는 내부수용감각이 추상개념의 체화에서 중요한 요인임을 의미한다. 실험2는 실험1의 결과를 뒷받침하였다. 내부수용감각 단어는 청각단어와는 구별되는 신경반응을 이끌어냈으며, 그것의 신경 원천은 내부수용감각관련 영역으로 추정되었다. 이 발견들은 추상개념의 체화가 내부수용감각을 포함한 감각운동영역의 활성화를 통해서 설명될 수 있음을 보여준다.

단어 감각 지수와 감각 특정 신경 반응에서 제1언어처리와 제2언어처리 사이에서 뚜렷한 차이가 발견되지 않았다. 이는 적은 환경적 접촉에도 불구하고 제2언어처리가 제1언어처리와 비슷하게 체화되어 있음을 보여준다. 하지만 두 실험은 또한 제2언어처리가

제1언어처리와 완전히 동일한 것은 아님을 드러내었다. 실험1에서 제2언어 학습자의 감각 경험은 그들의 모국어에 의해서 영향을 받아서, 몇몇 단어의 감각 지수 분포는 영어가 아닌 그들의 모국어인 한국어와 더 유사하였다. 실험2에서 제2언어처리는 시간-주파수 분석에서 감마 영역대의 지형도와 극성의 차이를 드러내었으며 소스추정분석에서 언어처리와 관련된 영역인 좌측두극(left temporal pole)의 추가적인 활성화를 보여주었다. 이런 결과는 제2언어처리의 체화가 제1언어처리에 비해 약화되었거나 언어적 경험에 더 의존함을 의미한다.

결론적으로 본 논문은 제2언어처리의 체화가 제1언어처리와 어떻게 다른지 내부수용감각과 추상개념과의 관계에 초점을 두어 탐구하였다. 두 실험을 통해서, 먼저 추상개념이 구체개념과 유사하게 감각운동 영역에 근거하고 있으며, 내부수용감각이 추상개념의 이해에 중요한 역할을 함이 드러났다. 또한, 본 연구의 발견은 제2언어처리가 제1언어처리와 비슷한 정도로 체화되어있으나, 보다 약한 정도의 체화와 언어적 경험에 대한 높은 의존의 특징도 가지고 있음을 보여주었다.

주요어: 체화된 인지, 추상 개념, 제2언어처리, EEG, 단어감각지수, 소스추정분석

학 번: 2017-32674