論文 / 著書情報
Article / Book Information

題目(和文)	
Title(English)	Language comprehension grounded in perceptual experiences: an investigation on text and speech perception
著者(和文)	Zhou Yizhen
Author(English)	Yizhen Zhou
出典(和文)	学位:博士(学術), 学位授与機関:東京工業大学, 報告番号:甲第12010号, 授与年月日:2021年3月26日, 学位の種別:課程博士, 審査員:金子 寛彦,小池 康晴,永井 岳大,長谷川 晶一,吉村 奈津江,渡邊 淳司
Citation(English)	Degree:Doctor (Academic), Conferring organization: Tokyo Institute of Technology, Report number:甲第12010号, Conferred date:2021/3/26, Degree Type:Course doctor, Examiner:,,,,,
学位種別(和文)	博士論文
Type(English)	Doctoral Thesis

LANGUAGE COMPREHENSION GROUNDED IN PERCEPTUAL EXPERIENCES: AN INVESTIGATION ON TEXT AND SPEECH PERCEPTION

Yizhen Zhou

Bachelor of Engineering (Hons), Electronic and Electrical Engineering, 2013
University of Bath

Master of Science, Information Processing, 2016
Tokyo Institute of Technology

Supervisor: Dr. Hirohiko Kaneko

External advisor: Dr. Junji Watanabe

Submitted in partial fulfilment of the requirements for the degree of

Doctor of Philosophy

Department of Information and Communications Engineering

School of Engineering

Tokyo Institute of Technology

ABSTRACT

There is now much evidence that language comprehension is grounded in sensory-motor systems to some extent. However, most previous studies have focused on the influence of dominate systems such as vision and action on individual word/sentence processing. The present research further investigated the grounding of language in three studies. The study 1 examined whether the influence of visual perception on spoken word processing can be observed at the sublexical (i.e., component parts of a word) level, and demonstrated that visual features induced a perceptual bias in geminate perception. The study 2 investigated the effect of touch, the relatively neglected sense, by using temperature stimulation. It was found that the processing of text was facilitated by congruent temperature stimulation. In addition, no effect of heartbeat vibration on narrative processing was found in study 3. Taken together, these results suggest comprehension of language involves perceptual experiences and is partially dependent on sensory-motor systems.

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my thesis advisor, Dr. Junji Watanabe for his continuous guidance throughout my PhD years and during the writing of the thesis. I also would like to thank my supervisor, Dr. Hirohiko Kaneko, for his support and help both in the research and university life. This thesis would not have been possible without their expertise and advice.

I would like to thank my committee members, Dr. Yasuharu Koike, Dr. Natsue Yoshimura, Dr. Shoichi Hasegawa, and Dr. Takehiro Nagai for reviewing this thesis and providing valuable comments.

My special thanks also go to Dr. Hsin-Ni Ho, Dr. Aiko Murata and Dr. Ryoko Mugitani, for their support and valuable suggestions for this work. I also would like to thank Mr. Yosuke Nakamura, his earlier work on geminate perception enabled me to build the foundation of the study 1.

I also want to thank my cat Mimi, for accompanying me while I was writing up the thesis.

Finally, I would like to thank my parents for their support and encouragement.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1
INTRODUCTION

1.1 OVERVIEW

Human beings communicate with each other using language in everyday life. It is a unique ability as no other species than humans learn and organize the associations between things in the external environment and arbitrary symbols in a system, which makes it possible to construct an unlimited number of sentences with a limited set of rules. Thus, it is important to try to understand the cognitive processes behind this universal human practice, i.e., how we use symbolic communication system to transmit meaning in the form of speech or text. Over the years, various theories have been proposed to explain how we process and understand meaning and concepts. For instance, the classical symbolic view argues that meaning of words arise from the combination of arbitrary symbols which are not related to the physical objects or external world, and they are independent of perception and action. In contrast, recent studies have proposed an alternative "embodied" view, which suggests that the concepts and the meaning of language are related to the experiences of perceiving the actual objects in the external world.

Since the 1980s, the idea that body matters to the cognition has been known as embodiment (Rosch & Lloyd, 1978; Johnson, 1987). From this perspective, the environment and the body are utilized in the cognitive system. In the brain's modal systems, the internal representations are implemented via simulations and interfaced with the external structures (Barsalou, 2010). As a result, higher level of human information processing such as language comprehension is grounded in organism's sensory and motor experiences (Barsalou, 2008; Wilson, 2002).

In the research of cognitive linguistics, many empirical studies have found evidence to support the idea that language comprehension is grounded in sensory experiences, and it is now well accepted that language comprehension is linked to perceptual experiences to some extent. However, most previous research has focused on individual word and sentence processing, whether the representations of visual features interact with language processing at the sublexical level has not been investigated. Moreover, relatively few studies have investigated the relation between touch and language comprehension, despite its importance to the body. In fact, the skin is the largest sensory organ in our body, and touch provides us the most fundamental and direct means to connect with the external world (Gallace & Spence, 2014). Therefore, it is necessary to examine the embodied hypothesis with the sense of touch.

In this thesis, a review of symbolic and embodied views of language comprehension is followed by the related empirical findings in order to set the stage for a series of experiments that assessed to what extent and how the perceptual experiences interact with language comprehension. In study 1, the influence of visual information on the perception of spoken words was examined. The study 2 investigated whether the relatively neglected sense, touch, interact with language processing. The study 3 further examined the influence of tactile processing on narrative comprehension. Although no effect of tactile information on narrative processing was found in study 3, the results from the two modalities (vision/touch) are compatible with the embodied perspective, particularly the weakly embodied view. In the following section, the research background on both amodal and embodied theories is reviewed. An outline of the thesis is provided at the end of this chapter.

1.2 LANGUAGE COMPREHENSION: MEANING AND CONCEPTS

The process of understanding meaning and concepts in language has been explained by classical symbolic view of amodal theory. The amodal theory of meaning and concepts proposes that understanding meaning and concepts involves the manipulation of abstract, arbitrary and amodal symbols (Collins & Loftus, 1975; Newell & Simon, 1976; Landauer & Dumais, 1997; Foltz, Kintsch, & Landauer, 1998; Landauer et al., 2006). This view argues words are considered to be external mediums mapping onto internal symbolic representations of word meaning (Weiskopf, 2010). The relationship between the symbolic representations and the real-world objects is arbitrary, and the meaning of the symbol is understood by its relation to other symbols. For instance, the word "chair" is understood because it is related to other semantic words such as "back", "leg" and "seat". In other words, the meaning of words is understood through the manipulation of symbols according to some specified rules (Weiskopf, 2010). The basic assumption underlying amodal theory is that perceptual inputs are transduced into representational symbols that describes these inputs amodally (Barsalou, 1999). Therefore, understanding the meaning of words does not involve the perception/action, nor does it recruit the sensory-motor system in the brain. Language comprehension is considered as being different from "lower-level" perceptual processes (Jirak et al., 2010).

In contrast, the embodied view suggests that the meaning and concepts are grounded in sensory-motor systems (Barsalou, 1999; Barsalou, 2008). That is, the meaning of a word is not only understood by the relationship between each word, but also by simulating the related perceptual experiences. For instance, the word "chair" can be perceived by touch (texture, weight), sight (color, size), and sound ("creak!"). It challenges the traditional symbolic view that core knowledge in cognition is represented in amodal structures and processed independently from sensory-motor system.

Over the past 25 years, the embodied view has risen to prominence in the study of language comprehension with accumulating behavioral evidence of interactions between language processing and sensory-motor processes (Barsalou, 2010; Glenberg, Witt, & Metcalfe, 2013; Meteyard et al., 2012; Meteyard & Vigliocco, 2008). To date, although there is still a debate between the classical symbolic view and the embodied view of cognitive/language processing, the majority opinion is that at least some of the mental representations need to be grounded in perception and action. Apparently, not all processing requires to be directly grounded in perceptual experiences. The conditions in which the comprehension system in language uses symbolic and embodied mental representations should be examined.

1.3 THE SYMBOLIC VIEW

The symbolic view of amodal theory is comprised of two principles according to Weiskopf (2010): Translation Hypothesis (TH) and Amodality Hypothesis (AH). The TH suggests that the main function of language is to translate between external (words or sentences) and internal representational medium (semantic or conceptual representations); the AH argues that the internal representations are amodal and distinct from the sensory-motor systems (Weiskopf, 2010). Ferdinand de Saussure, "the Father of Linguistics", has suggested that "language is a system of interdependent terms in which the value of each term results solely from the simultaneous presence of the others" (de Saussure, 1959). From this point of view, symbols are manipulated in mind as a high-level process in contrast to the activities that are operated as peripheral and "lower-level" process, i.e., perception and action. The low and high processes are seen as independent to each other.

In amodal theory, meaning and concepts are considered as independent from the body and the external world. They are represented propositionally in cognition, for instance, meaning is represented and processed through semantic networks or properties (Fodor, 1998; Phylysin, 1973). According to this framework, a process of transduction occurs when the perceptual experiences from the external world is getting processed in the brain. Such process results in frozen representations of the physical world: the link between the mental representations and the actual objects in the surrounding world is lost in the process. Therefore, the representations are just arbitrary and abstract symbols that have no specific modality features. For instance, the concept/meaning of the word "chair" is related to the propositional and amodal feature "it creaks", rather than related to the modal-featured acoustic feeling when hearing the sound of a chair creaking. In other words, language can be seen as a specific device used by human that obeying specific rules to manipulate arbitrary and abstract symbols. These symbols are not dependent on sensory-motor systems, as they are organized and operated stably under some particular linguistic rules.

Following this view, many statistical models of semantic memory have been elaborated to represent the meaning of words that extracted a large corpus of existing text by using statistical computations. These models assume that the knowledge of vocabulary is organized propositionally, and the meaning of a word is dependent on the relatedness with other words and the lexical cooccurrence. Examples include the Hyperspace Analogue to Language (HAL, Burgess & Lund, 1997) and the Latent Semantic Analysis (LSA, Landauer & Dumais, 1997). Both of the models represent the word meaning as vectors, and the meaning of words can be detected in matrices that contain the information of lexical cooccurrence. That is, the meaning of a word is considered as fixed and can be derived by the relationship between other words.

Although several experimental results are compatible with the statistical models of semantic words, for instance, they predict reaction times (RTs) in lexical priming experiments (e.g., Jones, Kintsch, & Mewhort, 2006; Hutchison et al., 2008), not every aspect of the LSA models work equally well. For example, one study found that participants can distinguish between LSA equivalent sentences depending on the perceptual features of the objects described in the sentences (Glenberg & Robertson, 2000). In this study, after seeing a sentence like "Marissa forgot to bring her pillow on her camping trip", participants judged the sentence "As a substitute for her pillow, she filled up an old sweater with leaves" more sensible and imaginable in sensibility and envisioning ratings than "she filled up an old sweater with water", regardless of the distance between the words in the matrices of semantic network (i.e., "water" and "leaves" are both far away from "pillow" in the matrices). In other words, people are able to distinguish between sensible and less sensible sentences quite easily, even when LSA models suggest that the sentences are equivalent (hence should have been difficult to distinguish). The results are against the view that language comprehension is based solely on fixed features and frequency constraints.

1.4 THE EMBODIED VIEW

In recent years, a rapidly growing body of research on "embodied cognition" argues that mental representations in cognition are not independent from the sensory-motor systems, but are rather grounded in the perceptual experiences (e.g., Barsalou, 1999, 2008, 2010; Glenberg & Robertson, 1999; Wilson, 2002; Zwaan & Kaschak, 2008). Therefore, human cognitive processes such as language comprehension, may be influenced by and associated with perceptual experiences and bodily states. The embodied view of language processing suggests the word meaning is not only understood by the relation of linguistic symbols, but also by simulating the related perceptual experiences. The fundamental idea of the embodied view is remarkably old and can be traced back to Plato's Cratylus: "there is a tight, natural and necessary relationship between the symbols of language and the real object that it represents."

1.5 A BRIEF HISTORY ON EMBODIED LANGUAGE COMPREHENSION

In the area of linguistics, the idea of embodiment has been used since the inception of cognitive linguistics. There are three distinct phases in the study of embodied language processing according to Bergen (2019). The first phase is the analytical phase that linguists attempted to look for evidence about how language might be embodied based on conceptual resources by the analysis of language. In this stage, the research did not focus on mechanisms, as it mainly investigated the online or developmental versions of embodied language processing. The second phase is called the process phase. The focus of the research was shifted to generate theoretical frameworks from empirical evidence. And the third phase is the functional phase of trying to determine the exact role of embodiment on particular aspects of the processing of language and cognition.

1.5.1 Analytical phase

Examples of the early work include (1) basic level categorization (e.g., Rosch et al., 1976), (2) polysemy (e.g., Lakoff, 1987), and (3) metaphor (Lakoff & Johnson, 1980, 1999).

(1) In free naming tasks of basic level categorization (Rosch et al. 1976), participants almost exclusively used basic category names (e.g., tree as opposed to the more specific pine or more general life form). These basic level terms were usually short and hence were easily learnt and accessed. Basic level terms seemed to be dependent on bodily interactions with the external word and shared common mental images.

(2) Polysemy means the words of multiple meanings. For instance, “hot” can be referred to both “heat” (temperature) and “spiciness” (taste). Therefore, for the word “hot” along, it may relate to both the heat and the spiciness, as there is similarity shared between the two experiences (Lakoff, 1987).

(3) Lakoff and Johnson (1999) have suggested the language is embodied in metaphor. For instance, the metaphor of affection and warmth (i.e., warm feels more “positive”). The experiences of warmth may refer to affection through metaphor, as there is a spontaneous relationship between physical warmth and affection.

1.5.2 Process phase

Examples of the theoretical models include Perceptual Symbol Systems (PSS, Barsalou, 1999) and Indexical Hypothesis (IH, Glenberg & Robertson, 2000; Glenberg & Kaschak, 2003). These models suggest language comprehension involves perceptual and motor representations. For example, understanding the word “lemon” involves modality-specific experiences and memories about how a lemon looks, smells, and tastes. The meaning and concepts in language are constructed in the brain by the activation of perceptual and motor systems to recreate the perceptual experiences of the object/scene described by the language. The common method for testing the theoretical hypothesis is usually to ask participants to perform a perceptual/motor task and a linguist task in some order. These two types of tasks are expected to interact with each other if language processing also recruited the sensory-motor systems in the brain. These tasks usually measure the participants' response times. For instance, one study asked participants to read sentences implying an object in a particular shape, such as a sunny-side up egg or a whole egg (see Figure 1-1, Zwaan et al., 2002). The participants then saw an image that showed the object in the same or a different shape as described by the just-read sentence, and they were required to judge whether the visual object had been mentioned in the previous sentence or not. The perceptual or motor effects are determined through facilitation or inhibition in participants' responses.

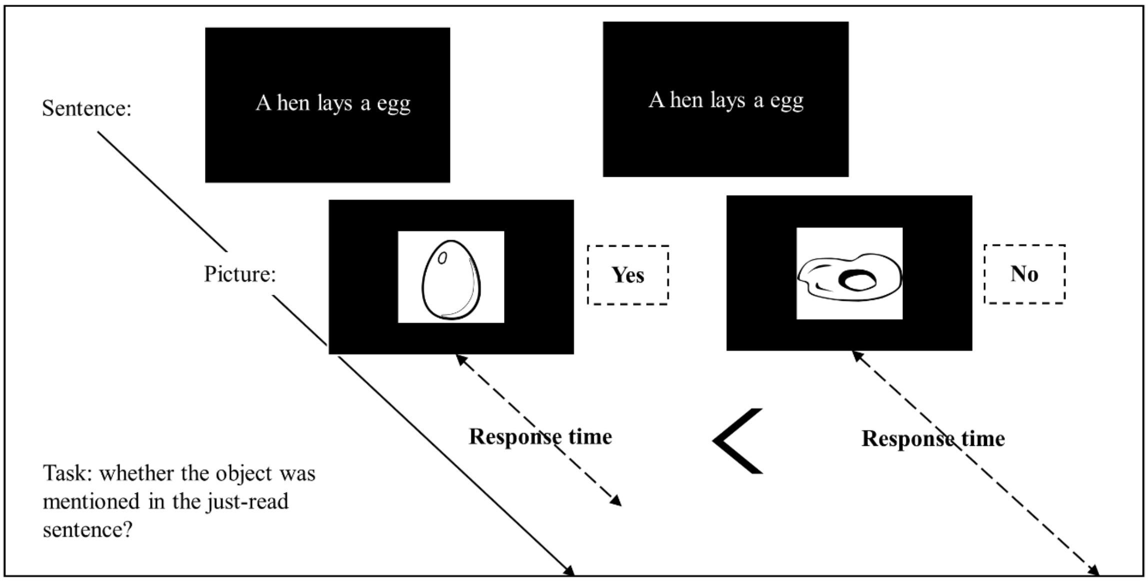


Figure 1-1. Sentence-picture matching task. Participants to decide whether the object was or was not mentioned in the sentence (based on the author's interpretation of Zwaan et al.'s work [2002]).

1.5.3 Functional phase

Brain imaging studies of embodied language have been conducted in brain damaged patients. For instance, it has been found that patients lost the access to words according to the damaged area of the brain. Particularly, patients had difficulties with nouns when they suffered the damage to left temporal cortex which was responsible for processing visual properties such as shape; and patient who damaged their left frontal cortex of motor control lost their access to verbs (Damasio & Tranel, 1993). In another study that fatigued a particular region of the brain (i.e., motor systems) by asking participants to perform a particular action by hundreds of times, the response times of participants were longer in a judgement task of sentences that describing the particular motions (Glenberg et al., 2008).

1.6 THE VIEW OF WEAK AND STRONG EMBODIMENT IN LANGUAGE COMPREHENSION

The theories of embodied language have been elaborated in different ways, and these theories (see Figure 1-2) might be categorized as weak or strong embodiment according to Meteyard and colleagues (2012).

	Label	Relationship to sensory-motor systems	Explanation of interactions	Neural correlates	Level of processing: word	sentence	narrative
FUSS (Vigliocco et al., 2004)	Weak embodiment	Partial dependence	Mediation	Distributed network proximal to sensory and motor systems	○	X	X
PSS (Barsalou, 1999)					○	○	X
LASS (Barsalou et al., 2008)					○	○	X
IEF (Zwaan & Ross, 2004)		Complete dependence	Modulation	Distributed network within sensory and motor systems	○	○	○
IH (Glenberg & Robertson, 1999, 2000)	Strong embodiment				○	○	○

Figure 1-2. Embodied language theories (weak to strong continuum). FUSS = Featural and Unitary Semantic Space; PSS = Perceptual Symbol Systems; LASS = Language and Situated Simulation; IEF = Immersed Experiencer Framework; IH = Indexical Hypothesis. ○ = discussed in the model X = not discussed in the model.

Basically, these theories suggest that representations of semantic content are at least partly grounded in sensory-motor systems. As the Figure 1-2 illustrates, the theories are divided as strongly embodied or weakly embodied. The theories can be grouped by the following parameters: (1) relationship to sensory-motor systems, (2) explanation of interactions, and (3) neural correlates (Meteyard et al., 2012). The parameter of "relationship to sensory-motor systems" refers to the degree of dependence that language processing has from sensory and motor information; the "explanation of interactions" defines how these theories explain the accumulated evidence that shows interactions between the processing of language and sensory-motor information; "neural correlates" refers to how brain processes such interactions. From a perspective of strong embodiment, the linguistic representations are completely dependent on sensory-motor systems, and the embodied effects are explained by direct modulation. The perceptual and conceptual processes are carried out within the primary sensory-motor systems. In contrast, weak embodiment view predicts a partial dependence of language processing on sensory-motor systems and explains the effect by mediation. Weak embodiment also suggests the integrated modal information is represented proximal to sensory-motor areas in the brain, and they are not processed by the same mechanism.

The Featural and Unitary Semantic Space hypothesis (FUSS, Vigliocco et al., 2004) is in line with the view of weak embodiment. It suggested that the meaning of word is grounded in conceptual featural representations, and the structure of concepts is organized by modality-specific (e.g., visual) feature types. The features are bound into semantic representations, and they are partly grounded in sensory-motor systems.

Another famous theory that can be classified as weak embodiment is the PSS (Barsalou, 1999). It covers not only the language processing, but also the human cognition. According to this theory, the link between the mental representation of an external object and the actual object is arbitrary (e.g., a “cat” is called “neko” in Japanese). However, there is an internal representation of the external object that extracts the perceptual experience perceived through the object. During the processing of the corresponding linguistic symbol, there is a reenactment of perceptual experiences induced by the perceptual symbol through simulation. That is, objects are represented in terms of perceptual symbols that grounded in perceptual experiences (Barsalou, 1999). Perceptual symbols are multimodal representations and can activate different sensory and motor information such as vision, audition, action, touch, smell and taste.

Barsalou et al. (2008) also proposed another theory named Language and Situated Simulation (LASS). It assumes that language comprehension is largely based on situated simulation. In the framework of LASS, there are two systems (i.e., simulation system and linguistic system). The simulation system captures the relation between properties in physical experiences, and the linguistic system is related to the linguistic features such as frequency of words or their relations to syntactic structures. Both systems are activated when language is processed. Although the representations of linguistic system peak first, they are relatively superficial.

According to the view of strong embodiment, the processing of sensory and motor information is part of routine semantic processing and this information is activated in primary cortex. Therefore, language comprehension is completely dependent on, and also directly modulated by sensory-motor systems. Theories such as Immersed Experiencer Framework (IEF, Zwaan, 2004) and IH (Glenberg & Robertson, 2000; Glenberg & Kaschak, 2003) both assume a "full" simulation of perceptual experiences through the modulation of activities in sensory-motor cortices.

In the theory of IEF, the perceptual experiences are proposed to be necessary in comprehension of language, as the sensory and motor information is certainly recruited in stimulation during the process of language comprehension. Therefore, comprehension is the reconstruction of the actual experience through integration and sequencing cued by the linguistic input (Zwaan, 2004).

The IH framework claims that language processing refers to situations, specifically. the "affordances" of a situation (Glenberg & Robertson, 2000; Glenberg & Kaschak, 2003; Kaschak & Glenberg, 2000). The affordances are based on the relation between objects and bodily abilities. The processing of language is hence linked to the preparation of action with the environment. The IH assumes words are indexed to real-world objects or to the perceptual symbols in PPS (Barsalou, 1999). The affordance recruited sensory and motor representations in a context dependent way. For example, a chair can afford humanlike bodies, but not large animals such as tigers to sit on it.

These embodied theories have predicted the influence of perceptual experiences at different level of language processing (i.e., at the level of word, sentence or narrative processes). The words, sentences, and narratives are assembled by letters, words, and sentences, respectively. The processing at the word level reflects the processes of matching orthographic and phonological information with the semantic properties of words drawn from a person's mental lexicon. The sentence processing involves higher level processes referring to assemble words into sentences. In the sentence processing, the information is not conveyed only at the word level but also at the level of syntactic structure. The processing of narrative is more complex and involves the process beyond the lexical and syntactic structure. Such process enables the reader/listener convey linguistic information in a larger context. In the narrative processing, the reader/listener acquires meaning at different levels that are not explicitly mentioned in the text or speech. In other words, the sentences in a narrative are integrated coherently to fill in the ambiguous part in the text/speech and construct representations of the character, states, scene and actions in the story. It has also been reported that the patterns elicited by brain activity at the word, sentence, narrative levels were different (Xu et al., 2005).

Among the embodied theories that mentioned above, two of them (i.e., IEF and IH) related to strong embodiment also concerning with the deeper level of language comprehension, i.e., narrative processing. In contrast, the narrative process is not disscussed in the weak embodiment models such as FUSS.

1.7 EMPIRICAL EVIDENCE OF EMBODIED LANGUAGE COMPREHENSION (TEXT AND SPEECH)

The past 25 years has provided ample empirical evidence showing that language comprehension interacts with perception and action in an online fashion. Most previous studies on embodiment and language have exclusively investigated the processing of individual words and sentences (Fischer & Zwaan, 2008). The embodied view of language comprehension predicts the activation in sensory-motor regions during linguistic tasks that involve the processing of semantic representations such as lexical decision, sentence-picture matching, and narrative judgment. It is also possible that the amount of sensory information recruited in the language processing is dependent on the nature of stimuli used in the task, and also on the level of language processing (Louwerse & Jeuniaux, 2010). The comprehension of individual word, sentence and narrative are likely to be different (e.g., Zwaan, 2004).

Many of the existing empirical findings are related to the dominant senses of vision, audition, and action. For example, studies have investigated the relation between spatial locations presented visually and word processing (Pecher et al., 2010; Šetić & Domijan, 2007). In their experiments, a congruency effect between word position and meaning was found. That is, participants responded faster to the words when the meaning of words and their spatial locations were matched (e.g., flying animals were processed faster when the words were presented at the top of the computer screen). Similarly, Estes, Verges, & Barsalou (2008) have demonstrated spatial congruency effect using context words. Each context word (e.g., cowboy) was presented following by an object word associated with either upper (e.g., hat) or lower (e.g., boot) locations. A target letter was then presented at the upper or lower locations of the screen, and the participants were required to identify the target as rapidly as possible. The responses were faster when the target appeared in the matching location than in the mismatching locations. Moreover, Meteyard et al. (2008) has demonstrated that responses in a lexical decision task were slower when the meaning of a motion word (e.g., rise) was incongruent with a visual motion pattern.

Evidence of the influence of visual perception has been extended beyond words to sentences for both speech and text. Sentence processing involves multiple-level processing including syntactic processing that goes beyond mere sequence. It has been found that the perception of speech is related to both visually presented motion and spatial representations. For instance, motions produced by visual stimuli influenced the action-related sentences: the concurrently viewed dynamic black-and-white stimuli affected the sensibility and grammaticality judgment of sentences describing motion in a particular direction (e.g., "The car approached you") presented auditorily (Kaschak et al., 2005). In the experiments, the motion produced by the visual stimulus was either in the same direction as the action described in the sentence (i.e., toward you) or in the opposite direction (i.e., away from you). Responses were faster to sentences presented simultaneously with a visual stimulus depicting motion in the opposite direction as the action described in the sentence. This result suggests that mechanisms of constructing simulations in language comprehension are also used in visual perception. Similar results were found in visual discrimination tasks of objects when hearing sentences contain nouns describing spatial representations (e.g., up or down) (Bergen et al., 2007).

For sentence processing in the form of text, its relation of visual perception of shape (Zwaan et al., 2002), orientation (Stanfield & Zwaan, 2001), distance of objects (Winter & Bergen, 2012), and size (de Koning et al., 2017) have been investigated using sentence-picture matching task. In the task, sentences and pictures of objects that were either mentioned in the sentence or not were displayed. Participants were required to decide whether the object was or was not mentioned in the sentence. The results indicate the responses were facilitated when the pictures were matched with the sentences. Similar facilitation effect of visual features was also found in processing of color information. For instance, Connell & Lynott (2009) asked participants to read sentences contained implicit information of color (e.g., sentences contained the word “tomato”) followed by a color-associated object word in typical, atypical or unrelated color. The color naming was facilitated when the word matched the implying color. There is also neuroimaging evidence that visual system is active when language conveying visual properties (Pulvermüller & Hauk, 2006; Simmons et al., 2007).

Interestingly, the studies investigating the relation between auditory perception and language comprehension have yielded different patterns in word and sentence processing. For example, although significant effects found in processing of visual objects when characteristic sounds (e.g., “meow” for “cat”, Iordanescu et al., 2008, 2010; Chen & Spence, 2011), hearing the same characteristic sounds did not facilitate the visual search of the names of objects (Iordanescu et al., 2008). However, there was an interference effect of hearing the motion sounds (e.g., the sounds of going “up”/ “down”, “toward”/ “away”) when judging the sensibility of target sentences: participants were faster to respond under the mismatching condition than under the matching condition (Kaschak et al., 2006). Similar interference effect was found with the target sentences presented in speech (Kaschak et al., 2006).

There is now ample evidence for the involvement of motor system specifically in speech perception. The role of motor system has long been recognized in the well-known motor theory of speech perception (Liberman et al., 1967; Liberman & Mattingly, 1985; Galantucci, Fowler & Turvey, 2006). The theory assumes that spoken words are perceived by identifying articulatory gestures and the coordinated actions of the vocal tract which they are pronounced rather than by identifying speech patterns. Several studies have demonstrated the activation of motor areas by passive speech perception (Rizzolatti & Craighero, 2004; Wilson & Iacoboni, 2006; Wilson et al., 2004). For example, Fadiga et al., (2002) has found that the activation of motor area was significantly different when listening to words and pseudowords containing tongue trilled or non-trilled sound: the muscle evoked potentials (MEPs) from tongue muscles were larger when listening to tongue trilled sound than non-trilled sound. That is, listening to linguistic stimuli at the word level produced the phoneme-specific activation of speech motor centers. This motor theory of speech perception seems in line with the embodied view of language comprehension.

Moreover, priming methodologies and lexical-decision task have been used in demonstrating priming effects of motor actions on speech processing. For example, Myung et al., (2006) have shown that similar motor action primed corresponding auditory words in a lexical decision task. In the experiment, participants' response times to target words (e.g., "typewriter") preceded by primes that shared manipulation features with the targets (e.g., "piano") were significantly shorter than an unrelated word (e.g., "blanket"). During the selection of stimuli, the possibility that such effect was induced by visual similarity of objects has been ruled out by selecting words rated as having high similarity in terms of manipulation but low visual similarity (such as the word pair "typewriter" and "piano"). The visual features of a typewriter and a piano are very different, while playing the piano and using a typewriter involve similar manual actions. Therefore, the words may mediate the primming effect by evoking action representations shared by the words automatically.

The embodied effects of textual words in action have also been examined in behavioral studies. For example, Gentilucci et al., (2000) has demonstrated a semantic effect on grasp aperture. In this experiment, an Italian word of either "grande" ("large") or "piccolo" ("small") was printed on an object of fixed size. The participants' movement directed towards to the object was affected by the concurrent semantic information: the "large" word evokes a larger maximum grasp aperture than the "small" word. Other than hand aperture, the influence of different types of hand postures afforded by objects has also been investigated. For instance, one study (Tucker & Ellis, 2004) asked participants to judge whether objects shown in pictures and object words were natural or man-made by using an input device that required either a power grip or a precision grip. The results indicate a compatibility effect on participants' responses: the responses of power grip and precision grip were faster to words and pictures implying objects requiring the same action than to word and pictures implying objects requiring a different action.

Taken together, these behavioral findings support the prediction that motor resonance occurs automatically during exposure to action related words presented visually and auditorily. There is also evidence from neuroimaging studies and studies using transcranial magnetic stimulation (TMS). For example, one famous study has demonstrated that the naming of tools selectively activated the left middle temporal gyrus and the left premotor area. These areas were also activated by action generation tasks and imagined hand movements (Martin et al., 1996). Stronger evidence showing motor resonance in language comprehension is provided by the findings that exposure to action and tool words evokes activation in motor areas of brain (Martin et al., 1996; Preissl et al., 1995; Pulvermüller, Lutzenberger & Preissl, 1999).

Pulvermüller (2005) has proposed the automaticity of motor-related activity for action words and a neural representation of action words (e.g., leg-related, arm-related and face-related words) in the motor and premotor cortices. This assertion is supported by a study using high density magnetoencephalography (MEG) (Pulvermüller, Shtyrov & Ilmoniemi, 2005). Participants took a distractor task while listening to action words related to the movement of the leg or face. The activations in premotor areas were found to be different for leg and face words, and these activations occurred within a short period of time (i.e., 170 ms) after onset of the words. Thus, it is unlikely that there is a contribution of strategic factors to the results. The cortical activation seems to be an early automatic process signified by the comprehension of action words. The time course of using neurophysiological methods including TMS and MEG can be found in a review (Hauk, Shtyrov, & Pulvermüller, 2008).

Since the embodied view of language comprehension hypothesizes a bidirectionality in language processing and perceptual experiences, the reversed process should also be found if the action words automatically evoked motor resonance. That is, recognition of these words should be facilitated by the stimulation of the related areas of motor and premotor cortices. The bidirectionality is support by a neurophysiological study that a facilitation effect was observed in a lexical decision task. In the task, participants responded to visually shown arm related words, leg related words or pseudowords by moving their lips. Their hand and leg areas in the left hemisphere were stimulated using magnetic pulses after the onset of words. The results indicate word recognition is facilitated by congruent motor action in related areas.

Behavioral studies have reached similar conclusion as in the neurophysiological research. For example, Lindemann et al., (2006) has demonstrated the lexical decisions were facilitated by congruent actions in a go/no-go paradigm. Participants' responses to the target words (i.e., eye or mouth) were facilitated following a congruent action (i.e., picking up a magnifying glass and bring it to the eye or to the mouth). This result suggests that semantic information associated with the goal of an action can be activated by preparing for the action, and hence the actions primed the understanding of action-associated words. However, it should be noted there were only two target words in the experiment. As a result, an expectancy effect may occur during the repetition.

Previous studies have also examined the motor resonance evoked by sentence processing. The embodied sentence comprehension has been demonstrated using sensibility judgment tasks of visually presented sentences. Glenberg and Kaschak (2002) have demonstrated an action sentence compatibility effect that action-compatible responses were faster than action-incompatible responses in judging the sensibility of target sentences such as "close the drawer". In this experiment, participants were required to response using buttons placed closer to or further away from their body. In the congruent conditions, participants needed to move their hand from the button close to the body, to the button away from the body, while judging the action sentence such as "close the drawer". Because in real-life scenarios, we usually move our hand away to close the drawer, the action performed by the participants was compatible with the action described in the sentence. The facilitation of movements has also been found using descriptive sentences with double object (e.g., "Andy delivered you the pizza"), dative constructions (e.g., "Andy delivered the pizza to you"), and even abstract transfer sentence such as "Liz told you a story". However, for abstract sentences such as telling someone a story, moving the hand away from the body is not an integral part. It is possible that the compatibility effect was evoked by the syntactic structure as such structure may have become associated with the movement of the hand toward or away from the body during early language acquisition (because these constructions are associated with verbs of manual transfer). The motor resonance evoked by sentence processing may be a complex phenomenon, as it may be evoked by individual words, the combinations of words, and even the linguistic construction. Moreover, the findings of Zwaan and Taylor (2006) indicate that motor resonance was not only occurred during lexical access but also during comprehension when information was integrated. A TMS study also assessed motor resonance in sentence comprehension and their results are consistent with the previous reports from behavioral research that the motor resonance occurred as the result of integration during comprehension (Glenberg et al., 2008)

There are also behavioral and neurophysiological studies investigating whether listening to action-related sentences modulates the activity of the motor system. For instance, Buccino et al. (2005) used both behavioral and TMS methods to investigate the influence of action-related sentences presented auditorily on the activity of the motor system. In the behavioral task, participants listened to sentences describing the actions of hand/foot and also the abstract sentences. The hand responses were slower during hand-related action sentences than during abstract sentences, and the responses were similar for foot-related action sentences. In the TMS experiment, participants listened to the same action sentences while their hand or leg motor area in the left hemisphere were stimulated by single-pulse TMS. Results showed that listening to hand-related sentences modulated MEPs recorded from hand muscles and listening to foot-related sentences modulated MEPs from foot muscles.

Table 1-1. Empirical evidence in embodied language comprehension

	Language comprehension (text and speech)		
	Word	Sentence	Narrative
Vision	△	○	△
Auditory	△	○	△
Touch	△	△	△

○: Well-established; △: Not fully investigated.

1.8 RESEARCH GAPS, OBJECTIVES AND SCOPE OF THE PRESENT RESEARCH

Overall, although empirical research has provided strong evidence that perceptual experiences such as vision or action influence comprehension of individual words and sentences, questions remain to be answered in the embodiment studies of language comprehension. Firstly, it was unclear whether the processing of visual features interact with language processing at the sublexical level (i.e., component parts of a word). In contrast to the lexical processing that corresponding to whole word forms, sublexical processing involves identifying constitute parts of words such as letters, phonemes and graphemes. Despite the well-established embodiment effect found in visual perception, the influence of visual features on phonological segments has not been investigated.

Moreover, relatively few studies have investigated the effect of a very important sense, touch, compared with the number of studies investigating relation between visual/auditory perception and language comprehension (see Table 1-1). Touch has been treated as one of the traditional five-senses aside from vision, audition, taste and smell (Sorabji, 1971). It is an important sense as it provides the most fundamental and direct means of contact with the external stimuli/objects to us (Gallace & Spence, 2014). We are able to acquire touch information and manipulate the external objects by touch ever since we were infants. There are receptors embedded in the skin and muscles to integrate tactile sensations, and hence induce the sense of touch from the activity of distinct neural systems responding to movement, pressure, temperature and the haptic properties of stimuli (Lederman & Klatzky, 2009). The perception of touch can be integrated with various sensory signals such as visual and auditory signals to induce meaningful perceptual experiences. For example, tactile acuity can be enhanced by visual processing (Kennett, Taylor-Clarke, & Haggard, 2001; Zhou & Watanabe, 2020).

According to the embodied view, the perception of touch should also be involved in language comprehension. However, relatively less was known about how language interacts with touch perception. Touch seems to be less described in language that the words used to represent the perceptual experiences of touch are fewer than to vision or audition (Levinson & Majid, 2014; Winter, Perlman, & Majid, 2018). Moreover, touch may also be difficult to define precisely as it has many “sub-senses” such as temperature or vibration (de Vignemont & Massin, 2015). Particularly in the research related to the language processing and temperature, an implicit association between emotional and thermal concepts has been revealed by text and temperature stimulation (Bergman et al., 2014). However, such result may imply an association between social interactions/experiences and physical temperature (i.e., warm is considered as more “positive” compared with cold) instead of reflecting the grounding of language processing on perceptual senses. That is, whether a direct link exists between temperature and processing of corresponding text/speech has not been investigated. Another example of showing association between touch (specifically the vibro-tactile sensing) and language processing is that aero-vibration influenced speech perception at the sublexical level: Gick and Derrick (2009) have demonstrated that the perception of consonants was influenced by inaudible air puffs stimulated on participants’ skin. Regardless, the research of embodied effect of touch was insufficient.

Another problem within embodiment studies on language comprehension is that the research focus has mostly been on individual words and sentences. Typically, decontextualized words and sentences have been used as stimuli in psychological experiments on language. This poses some problems because results obtained in brain and behavioral studies cannot be simply extrapolated to narrative/discourse comprehension. Compared with the single word or sentence comprehension, narrative/discourse comprehension comes closer to mental imagery and there are extensive mental representations generated during narrative processing (Zwaan & Radvansky, 1998). Moreover, we rarely understand words or sentences presented in isolation in real life. Investigating the influence of perceptual experiences on narrative processing may indicate to what extent and how the embodied representations are involved in language comprehension.

To overcome these research gaps, the present research investigated to what extent (at the levels of the word [sublexcial and lexical] and the contextualized sentence) and how the embodied representations interact with language processing, rather than trying to find evidence for or against a particular theoretical framework (e.g., PSS). This thesis further investigated the influence of perceptual experiences on language comprehension. First, most previous research has focused on word and sentence processing. In contrast, the current thesis demonstrated that spoken word processing may also related to visual objects at the sublexical level. Second, whereas most previous studies have focused on visual and auditory perception, the current thesis investigated the effect of a very important but less studied sense, touch. The embodiment effects found in study 1 and 2 from two different modalities (vision/touch) on speech/text processing confirm the multimodal nature of language and extend the embodied theory with new empirical evidence from an important but relatively less studied perceptual experience, i.e., touch. In the present study, two “sub-senses” of touch (i.e., temperature and vibration) have been investigated. The influence of vibration was investigated in the form of another type of the bodily sense, interoception, although technically it is not a component of touch. Although no effect of interoception information provided through touch was found on narrative processing in study 3, the results of sensory experience of touch indicated there is a direct link between tactile perception and text comprehension.

Taken together, the results of the present research suggest that the relationship between language comprehension and sensory-motor system is partially dependent, and it is unlikely that perceptual and conceptual processes are carried out by the same mechanism. Instead of a direct modulation predicted by strong embodiment view, the embodied effects may be explained by mediation, which is consistent with the view of weak embodiment.

1.9 OUTLINES OF THE THESIS

The thesis is organized in the following manner: Chapter. 1 Introduction, Chapter. 2 Study 1, Chapter. 3 Study 2, Chapter 4. Study 3, and Chapter 5. Conclusion. In Chapter 1, two views of language comprehension are introduced: the amodal theory and embodied theory. The Chapter 2 describes the experiment of the influence of visual information (illustrations) on spoken word categorization. The Chapter 3 describes the experiment of the influence of touch on text word categorization. In Chapter 4, the experiment investigating the influence of touch (heartbeat vibration/teardrop) on narrative processing of news stories is described. Finally, the Chapter 5 summarizes the conclusion of the entire thesis and suggests the directions of future research.

CHAPTER 2

STUDY 1: PRIOR VISUAL INFORMATION (ILLUSTRATIONS) INFLUENCED SPOKEN WORD CATEGORIZATION

2.1 Method
2.2 Results
2.3 Discussion

Among human sensory experiences, a large part is visual. Vision not only provides us the precise spatial and temporal information about the external world, it also seems to dominate lives of sighted individuals. It is one of the most intensively investigated human senses, and previous empirical research of embodiment has also focused on the sense of sight. There is now sufficient evidence that visual perception matters to our ability to understand language about comprehension. However, most previous studies have focused on the influence of visual information on individual word or sentence processing (e.g., Bergen et al., 2007; Estes, Verges, & Barsalou, 2008; Kaschak et al., 2005; Meteyard et al., 2008; Pecher et al., 2010; Šetić & Domijan, 2007). To determine to what extent the language processing is embodied in visual perception, it is necessary to investigate how the visual information influence word processing at the sublexical level.

In study 1, the influence of visual information on perception of spoken words was examined. As mentioned in Chapter 1, the motor theory of speech perception has suggested the active participation of motor system in speech perception (e.g., Liberman et al., 1967; Liberman & Mattingly, 1985). According to the motor theory, people perceive the sounds of speech by identifying how those sounds are produced. The key assumption that the motor system is recruited for speech perception has led to new research of cognition. In particular, it has inspired later neuroscientific studies of mirror neuron system. The discovery of mirror neurons suggests that motor system is involved in perception as these neurons are responsible for both of the observation and execution of identical muscular efforts (Gallese et al., 1996; Rizzolatti & Craighero, 2004; Rizzolatti, Fogassi, & Gallese, 2001). Particularly for speech perception, there is sufficient evidence that perceiving speech involves activity of motor cortex (D'Ausilio et al., 2009; Fadiga et al., 2002; Möttönen & Watkins, 2009; Pulvermüller et al., 2006; Schomers et al., 2015; Tettamanti et al., 2005; Watkins, Strafella, & Paus, 2003). Among these studies, it has been determined that such specific perception-action links in speech perception can be found at the level of speech sounds (phonemes, e.g., Möttönen & Watkins, 2009), words (e.g., Schomers et al., 2015) and sentences (e.g., Tettamanti et al., 2005). For instance, Möttönen and Watkins (2009) have shown that categorical perception of artificial acoustic continuum involving lib movement (e.g., ba/–/da/) was impaired by repetitive transcranial magnetic stimulation (rTMS) that disrupted the lip representation in the left primary motor cortex. In contrast, speech sounds that do not require lip movement during articulation (e.g., /ka/-/ga/) did not produce the same effect. In addition to the demonstrations of cortical motor activation in conjunction with speech perception, one study has demonstrated somatosensory input produced by stretching the facial skin influenced the perception of speech (Ito, Tiede, & Ostry, 2009). The perceptual boundary of the stimulus continuum between the words "head" and "had" was influenced by patterns of facial skin deformation that was similar to the production of the words "head" and "had". Therefore, speech processing is likely to be grounded in the perception and physical activities. The perceptual experiences may induce a perceptual bias even at the sublexcial level.

Previous research has already revealed the influence of visual information in the form of written words on spoken word processing. For instance, visual presented words can produce a priming effect on identical spoken words (Koudier & Dupoux, 2001; Grainger et al., 2003). From a perspective of embodied language, the illustrations or photographs may induce a similar effect if the visual features are simulated during language comprehension, and they can serve as alternative representations of the identical written words. One study has demonstrated a priming effect on spoken words in an identification task using photographs when the context of photographs and target stimuli were identical (Shigeno, 2017). The visual information perceived consciously in the form of illustration may influence the perception of spoken words similarly as the identical written words do. To investigate whether such influence exists or not, the experiment of study 1 examined the influence of illustrations specifically on perceptual categorization of Japanese geminates using a categorization task. In contrast to the previous studies that typically using the tasks of measuring reaction time (RT), the categorization task directly reflects a perceptual change occurred during the processing of the target stimuli.

Many languages including Japanese exhibit a durational contrast between singleton and geminate consonants (Ladefoged & Maddieson, 1996). The geminate consonants (known as “sokuon” in Japanese) are considerably longer than singleton consonants. Previous studies have examined the acoustic correlates of singleton/geminate consonants and they found that the primary correlate is the duration (e.g., Idemaru & Guion, 2008). There are two types of geminates in Japanese language: silent (e.g., /tt/, /kk/) and fricative geminates (e.g., /ss/). It has been found that native Japanese speakers have a specific tendency when listening to geminates: they share an abstract representation of geminates similar to a silence (Sadakata et al., 2012; 2014). Japanese phonology and orthography may induce such tendency, and the linguistic knowledge of native speakers is more heavily influenced by orthography compared with non-natives (Nagano & Shimada, 2014). Japanese geminates are moraic because Japanese is a more-time language. The closure duration in a silent geminate or sustained frication in a fricative geminate aligns with the duration of a mora (i.e., a perceptual unit of timing) (Vance, 1987; Han, 1994). For instance, there are four moras in a word /batta:/ (“batter”) that having a geminate /tt/, while the word /bata:/ is a three-mora word having no geminate. In Japanese, the geminates are represented by “ッ” (small “tsu”) regardless of the acoustical types, while the two types differ in occurrence. The silent geminates appeared to occur more frequently than fricative geminates. In the 5th edition of Iwanami Japanese dictionary, the percentage of silent geminates was much greater than fricative geminate (76 percent versus 24 percent among 2,476 geminate consonants) (Takahashi, 1998). Therefore, the present study focuses on the perception of silent geminates (referred to as geminates in the rest of this thesis).

As for the perception of geminates, the most important perceptual cue distinguish between singletons and geminates used by Japanese speakers is the closure duration (Kawahara, 2015). A target word is more likely to be perceived as geminate word with a longer closure duration (Amano & Hirata, 2019; Kinston, Kawahara, Chambless, 2009; Oba, Braun & Handke, 2005). In the perceptual studies of geminates, the perceptual boundary between singleton and geminates have usually been determined. The perpetual boundary is the durational value corresponding to categorical shift in perception of singleton/geminate. As a result, the changes in perceptual boundary induced by external stimuli can indicate the influence of such stimuli on the perception of geminates. For example, both the preceding vowel duration (Kingston, Kawahara, & Chambless, 2009) and the speech rate at sentence level (Hirata & Whiton, 2005) influenced geminate perception.

Similarly, the present study investigated the influence of illustrations by determining and comparing the changes in the perceptual boundary of target words preceded by primes. In this study, a minimal pair of words (the singleton word /bata:/ ["butter"] and the geminate word /batta:/ ["batter"]) that has the same phonemes but different closure duration was used. The experiment assessed the influence of perception of geminate /tt/ and the related singleton /t/ by observing the perceptual bias induced by different primes. The perceptual bias of consonants may reflect an embodiment effect at the sublexical level.

The embodied view predicted that both the written words and illustrations would influence the perceptual categorization of the spoken words, as understanding the meaning of words involves activating modality-specific representations or processes. The concept of the word "butter," for example, is formed by the associated memory images of visual and taste perception. Therefore, an illustration that matches the implied linguistic content may activate semantic processing and hence affect the spoken word processing.

2.1 METHOD

Participants. Twenty-four paid participants (17 females) aged between 20 and 40 years participated in the experiment. All participants were native Japanese speakers. They had normal or corrected-to-normal vision, and no hearing impairments. This experiment was approved by the Ethics Committee of NTT Communication Science Laboratories and performed in accordance with the ethical standards set down in the 2013 Declaration of Helsinki. Written informed consent was obtained from all participants in advance.

Apparatus. The experiment was conducted in NTT Communication Science Laboratories, Japan. A personal computer (Mac Pro) running Matlab 8.5 with PsychToolbox 3 was used to control the presentation of stimulus and collect the participants' responses. The participants were seated in front of a 13.3' Mac Pro computer display with a resolution of 2560 x 1600 pixels and a refresh rate of 60 Hz. The built-in speaker of the Mac Pro presented all of the auditory stimuli, and the average sound levels were normalized. The participants responded to target stimuli by using the computer keyboard ('←' and '→' keys). The experiment was conducted in a dark and quiet room.

Stimuli. The sound stimuli (/bata:/, /batta:/) were recorded by a professional native Japanese narrator in a soundproof room at a sampling frequency of 44.1 kHz. The intervals /ba/ and /ta:/ were extracted from the base stimulus of singleton /bata:/, and then the silent intervals were inserted between each of the /ba/ and the following /ta:/. The closure durations between /ba/ and /ta:/ were edited to range from 130 to 210 ms in 10-ms steps to create a /ba/ - /ta:/ continuum of nine auditory stimuli (see Figure 2-1 for the waveforms). Before the start of experiment, whether the native Japanese listeners were able to perceive the pair of stimuli with the closure duration of 130 ms and 210 ms as singleton (/bata:/) and as geminate (/batta:/), respectively, has been tested. Twenty-seven native Japanese speakers including 22 participants in the experiment were asked to identify the stimuli without presenting any prime. Each stimulus was randomly repeated 25 times. The proportion of judging 130 ms and 210 ms as geminate was 0.296 and 99.1 percent, respectively. This means that participants could surely perceive the stimulus containing the closure duration of 130 ms as /bata:/, and 210 ms as /batta:/. For prime stimuli, the spoken word primes were hence either /bata:/ (closure duration of 130 ms between /ba/ and /ta:/) or /batta:/ (210-ms closure duration). Figure 2-2 shows the written word and illustration primes used in the experiment. The size of the illustrations used in the experiment was 710 x 709 pixels. All nine auditory stimuli were used as the target stimuli.

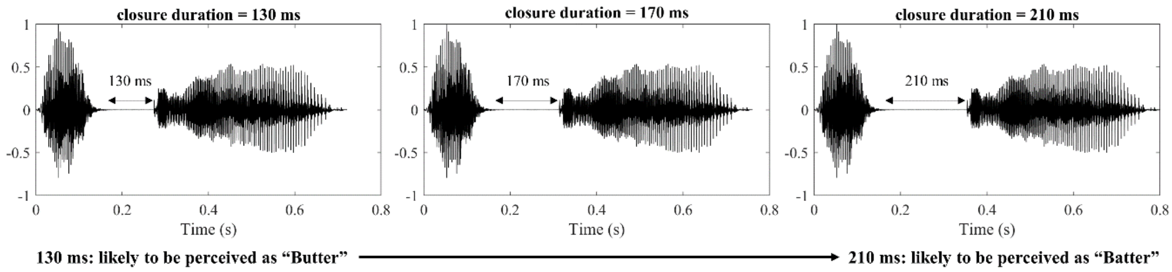


Figure 2-1. Waveforms of auditory targets /bata:/ and /batta:/.

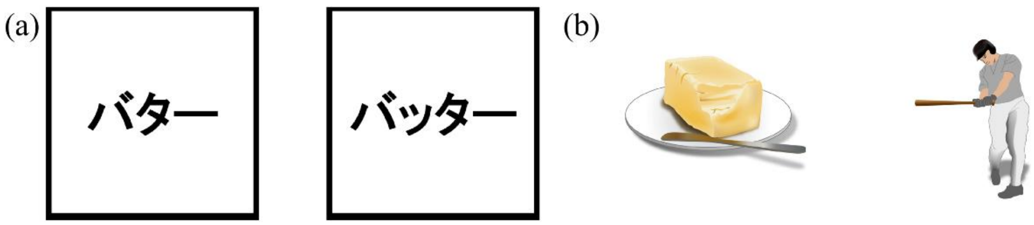


Figure 2-2. Prime stimuli: (a) Written word primes: /bata:/ and /batta:/. (b) Illustration primes: /bata:/ and /batta:/.

Procedure. In the experiment, a categorization task was conducted that participants identified the target auditory stimuli (one of nine closure durations) as either singleton word /bata:/ or geminate word /batta:/. The participants were seated in front of the computer, and instructed to fixate the center of the screen and respond to the target stimuli as accurately as possible by pressing one of two response keys (the 'singleton' and 'geminate' keys) on the computer keyboard. Three experimental conditions were created:

a spoken word prime condition, written word prime condition, and (3) illustration prime condition. Each experimental condition involved 450 trials (see Figure 2-3 for the trial structure), and these trials were separated into five blocks (90 trials per block). In each block, all the prime-target pairs (2 categories (singleton or geminate) of primes x 9 types of targets) were randomly repeated five times. Therefore, to complete all three conditions, each participant took part in a total of 1350 trials (90 trials x 5 blocks x 3 conditions). Under the first condition, the spoken word primes (either /bata:/ with a 130-ms closure duration or /batta:/ with a 210-ms closure duration) were presented prior to the presentation of the target stimuli. The durations of the word /bata:/ and /batta:/ were 660 and 740 ms, respectively. Each trial began with a white noise sound (in a 20 Hz to 20 kHz frequency range) to remove the effect of the previous trial. Under the second condition, the prime was the written word /bata:/ or /batta:/. The written word primes were presented randomly for 700 ms (i.e., the average duration of the spoken word primes) prior to the presentation of the target stimuli. To eliminate the effect of the previous trial, each trial began with a noise image (gray mosaic). Under the third condition, the primes were changed to the illustration of /bata:/ (butter) or /batta:/ (batter). The procedure was exactly the same as for the written word primes. The order in which each condition (5 blocks in succession) was presented was counterbalanced across participants. Participants took a 25-minute rest between each condition.

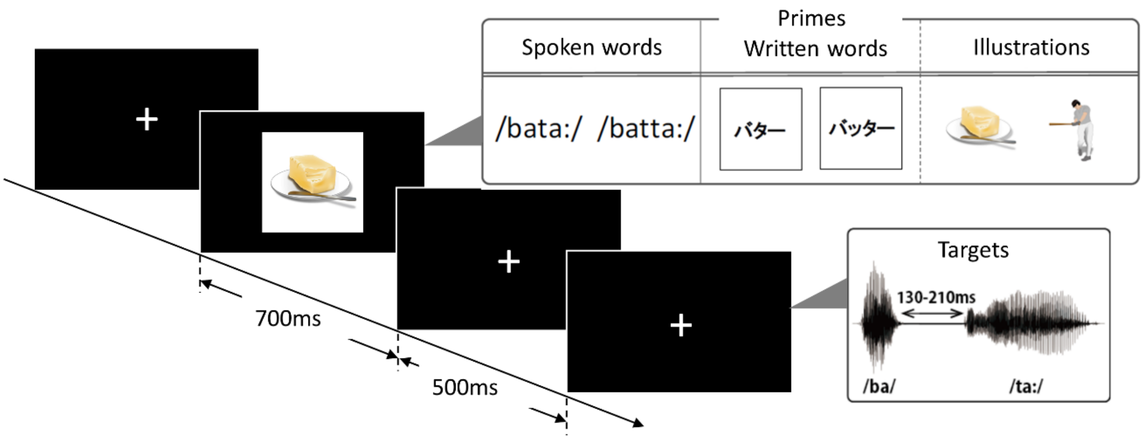


Figure 2-3. A schematic representation of the events on each trial.

Data analysis. The participants' judgments of the target stimuli were used in the analysis. The proportions of the target stimuli judged to be the geminate word /batta:/ were calculated for each prime category (singleton and geminate) to fit a psychometric function. Each psychometric function was fitted by a cumulative logistic distribution function to the data with the maximum likelihood method to estimate the 50% response point, i.e., the point of subjective equality (PSE). Therefore, the closure duration at this PSE is the perceptual boundary of the singleton word /bata:/ and the geminate word /batta:/. Figure 2-4 shows a representative example of the priming effect. The solid and dashed lines in the figure are the psychometric functions for the spoken-word prime condition obtained for participant M.Y in the experiment.

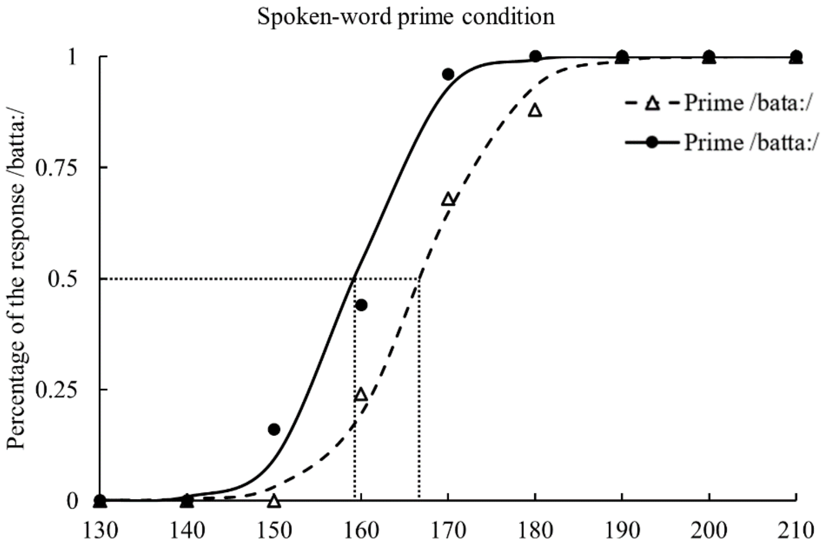


Figure 2-4. Psychometric functions for the spoken-word prime condition obtained for participant M.Y. Circles and triangles show the raw data; solid and dashed lines show the fitted psychometric curves.

In addition, a linear mixed effect analysis on the participants' judgments of the target stimuli was conducted. The lmer function in the R package lmerTest (Kuznetsova, Brockhoff, & Christensen, 2017) was used to build the linear mixed effect model. The percentage of the geminate response was used as the dependent variable. The independent variables included prime category (singleton or geminate), modality (auditory or visual) and the interaction between prime category and modality. The target stimuli and participant were included as random intercepts. The significance of the individual effect was tested by likelihood ratio tests comparing the full model with a reduced model without the effect.

2.2 RESULTS

The group means of the closure durations at the PSE for the three conditions are shown in Figure 2-5. Using the obtained closure duration boundary as the dependent variable, a two-way repeated measures analysis of variance (ANOVA) using SPSS Statistics software was conducted. The within participants factors were the category of primes (singleton word /bata:/; geminate word /batta:/) and conditions (spoken word primes, written word primes, illustration primes). The main effect of the prime category was significant [F (1, 23) = 20.622, p = .000, $\eta 2$ = .473]. The main effect was also observed between conditions [F (1.445, 33.228) = 7.528, p = .005, $\eta 2$ = .247]. No significant interaction was observed between prime category and condition [F (1.429, 32.876) = 3.064, p = .076, $\eta 2$ = .118]. The results indicate that when the prime stimuli of /batta:/ were presented, the participants were more likely to judge the target stimulus as the geminate word /batta:/; and vice versa when /bata:/ was presented. Bonferroni's method for adjusted pairwise comparison revealed that the perceptual boundary of the auditory (spoken word) prime condition was significantly larger than that of visual (written word/illustration) conditions [for the auditory and visual (written word) prime condition, p = .024; for the auditory and visual (illustration) prime condition, p = .021; for the visual (written word) and visual (illustration) prime condition, p = 1.000]. The ANOVA results were further supported by a Wilcoxon signed-rank test (geminate primes induced a smaller perceptual boundary than singleton primes among 18, 19 and 15 out of

24 participants under the spoken-word prime, written-word prime and illustration prime condition, respectively. For the spoken-word prime condition: Z = -3.021, p = .003, η2 = .617; for the written-word prime condition: Z = -2.623, p = .008, η2 = .535; for the illustration condition: Z = -2.721, p = .007, η2 = .555). In other words, a target stimulus with the same closure duration was more likely to be judged as the singleton word /bata:/ in the auditory condition than in the visual condition, as the perceptual boundary in the auditory condition was larger than that for the visual condition.

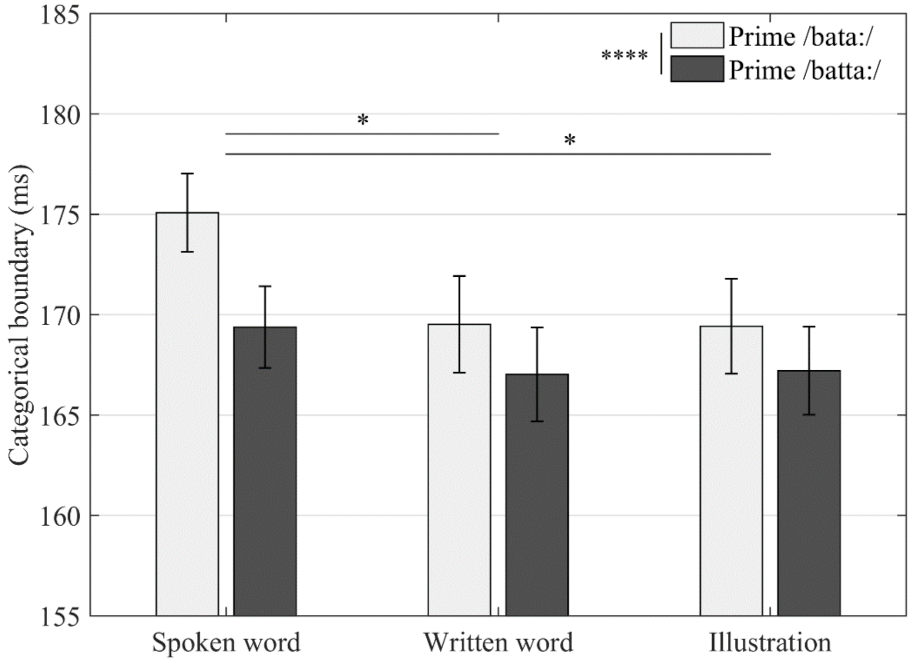


Figure 2-5. Results indicating the group means of the closure durations at PSE for the three conditions. The prime stimuli in the two categories are indicated by the bar color, with light and dark bars representing the prime stimuli of singleton /bata:/ and geminate /batta:/, respectively. The error bars show the standard error of the mean. * indicates the statistical significance of p <.05, and **** indicate the statistical significance of p <.0001 (three-way repeated measures ANOVA). N = 24.

The linear mixed model analysis was conducted to assess the fixed effects of prime category and modality. A full model (with fixed effects of prime category and modality) was constructed, and the reduced models was compared against it to specifically examine the effects of the prime category and modality. Likelihood ratio tests confirmed that reduced model lacking the effect of prime category significantly differed from the full model ($\chi2 = 25.688$, $p = .000$). The effect of modality also had a clear impact on responses of geminates ($\chi2 = 24.813$, $p = .000$). There was no significant interaction between geminate category and modality (likelihood ratio test comparing the full model and the reduced model without the interaction: $\chi2 = 3.435$, $p = .064$). In the experiment, although the participants were more likely to perceive the targets as geminates when geminate primes were presented both in auditory and visual modality, a target was more likely to be perceived as a geminate word if the word primes were presented in visual modality. The full model results are summarized in Table 2-1.

Table 2-1. The results from the linear mixed effect model

Fixed effects	B	Std. Error	T	p
(Intercept)	.50704	.13248	3.827	.00367
Modality	.02519	.01274	1.976	.04834
Prime category	-.06259	.01472	-4.254	.00002
Modality: Prime category	.03343	.03343	1.855	.06387
Random effects	**SL**		**Variance**	**Std. Dev.**
Participant ID	(Intercept)		.00889	.09427
Target stimuli	(Intercept)		.15366	.39199
Residual			.02339	.15292
Model				
AIC	BIC	logLik	deviance	df. resid
-1043.1	-1006.9	528.6	-1057.1	1289

2.3 DISCUSSION

The present experiment examined the influence of auditory and visual primes on Japanese geminate perception. In the experiment, a perceptual categorization task for

Japanese singleton and geminate words was used in a priming paradigm. The changes in the perceptual boundary of the singleton/geminate words were estimated. The results indicate that both auditory and visual primes, even the illustrations, induced a perceptual bias in geminate perception. The primes of singleton word led to an increase in the perceptual boundary. In contrast, the perpetual boundary was shortened by the geminate word primes. In other words, the target stimuli were more likely to be perceived as geminate words when participants heard a geminate-featured prime. Moreover, the influence of prior auditory and visual information on geminate perception may differ. The perceptual bias induced by auditory primes was different compared with written word and illustration primes.

Previous research has demonstrated a multimodal nature of speech perception. For instance, it has been well established that visual information regarding facial detail influences speech perception and helps listeners distinguish sounds degraded with noise (Buchwald, Winters, & Pisoni, 2009; Erber; 1969; Summerfield, 1979). The famous McGurk effect and its subsequent studies have shown that the perception of internal phonetic category is also sensitive to visual information (Brancazio, Miller & Paré, 2003; McGurk & MacDonald,1976; Tippanna, 2014). Moreover, the effects of somatosensory information on speech perception have been reported (Gick & Derrick, 2009; Ito, Tieda, & Ostry, 2009).

The multisensory nature of the speech has also been investigated in cross-modal studies of speech perception. Previous studies have typically the lexical decision task in a priming paradigm. In the priming paradigm, a prime and a target are presented in close temporal succession and participants perform a task such as lexical decision. It has been revealed that visual words influenced spoken word processing in repetition priming (Grainger et al., 2003; Grainger & Ferrand, 1994; Kouider & Dupoux, 2001). In the lexical decision task, participants judge whether a target stimulus is a word or not, and the RTs and error rates are measured. The task is specifically effective for testing the particular process engaged in evaluating the lexical status of stimuli. However, given the assumption that visual information plays a role during spoken word processing, one would expect priming effects to be observed even in perceptual tasks such as categorization that reflects a direct perceptual change in target processing other than measures of RT.

In the present study, a significant bias was observed in perception of closure duration for auditory primes. It is consistent with previous research showing that responses to the target words are facilitated when the auditory primes and targets were phonologically similar (Goldinger et al., 1992; Radeau, Morais, & Segui, 1995; Slowiazcek, Nausbaum, & Pisoni, 1987). Moreover, according to the view of embodied language processing, the non-orthographic visual information (e.g., illustrations or photographs implying the visual features) would also induce the perceptual bias because the illustrations may serve as alternative representations of the identical words. In contrast to the previous studies investigating the influence of visual information on speech perception, the present study used a perceptual task to examine the influence of illustrations on the perception of spoken words. The results show that illustrations also influence geminate perception when their meanings are matched. It provides evidence that illustration may evoke the related phonological representation, and hence listeners use it to inform the perception of subsequent auditory targets. In addition, the present study not only revealed the direct influence of illustrations on the perception of spoken words just as photographs do (Shigeno, 2017), but also demonstrated illustration primes generated significant bias just as written word primes did. The phonological representations may be activated by written words and illustrations in a similar way, and their influence on spoken word perception was a result of a match in phonological representations between primes and targets. Such influence can be observed at the level of sublexcial processes. These results are consistent with the perspective of embodied language processing that the sensory-motor system is involved in the perception of speech.

To summarize, the present experiment was designed to determine the influence of prior auditory and visual information on geminate perception. As observed here, the closure duration presented in the form of auditory signal, visual orthography (written word) and even non-orthographic visual information (illustration), can guide boundary shift for Japanese geminate perception. In addition, the results indicate that the phonological information presented in auditory and visual modalities are processed differently: the prelexical phonological information is activated when mapping auditory signal and mental lexicon but the visual orthography may only activate lexical phonological information; the auditory information is critical for inducing a perceptual change in geminates, and visual information appears to have a different impact that lexicality is required for information presented in the visual modality.

CHAPTER 3

STUDY 2: TOUCH INFORMATION (TEMPERATURE) INFLUENCED TEXT WORD CATEGORIZATION

There is now a great deal of empirical evidence supporting the embodiment view of language that language comprehension involves sensorimotor representations (Meteyard & Vigliocco, 2008). Alongside the numerous behavioral findings of interactions between language processing and sensory-motor systems, it seems reasonable to reject the idea that meaning is only composed of arbitrary, abstract and amodal symbols. To date, the latent majority view of language comprehension considers that at least some of mental representations is grounded in perceptual experiences. However, the extent to which language is grounded in perception and action has not been fully understood. There is a limitation in previous empirical research of embodied language. That is, most studies have focused on the effects of vision, audition and action, while other senses such as touch has been neglected in embodiment research of language processing. From a perspective of embodied language processing, touch (as one of the senses that contribute to perceptual experiences) should also be recruited for language comprehension.

Among all the human senses, touch is a unique sense and one of the necessary sources for sensory information processing. The sense of touch is essential for our everyday experiences. It contributes to the processing of material characteristics of surface and object. Moreover, it may even set the boundaries of the self to separate us from the surrounding environment (Gallace & Spence, 2014). Touch is often thought to be the sense that cannot be easily fooled or deceived. We would usually touch something when we try to convince its actual existence. Although touch is very important for our everyday lives, historically little was known about touch in comparison to vision and audition. Therefore, it might be difficult to predict how language processing would interact with touch. Moreover, it is relatively difficult to characterize the touch modality because it has many "sub-senses" such as temperature, vibration, tactile sensing of roughness or hardness, and so on.

Although touch may be difficult to define precisely, there is no doubt that temperature is one of the distinct components. Temperature perception is crucial for humans in daily activities, as it reflects our physiological conditions and we need to maintain desirable temperatures to avoid harm. Temperature is also indispensable for identifying a particular material or object in external world, as the elimination of thermal cues (heating various materials to make their temperature close to the skin temperature) results in impaired material discrimination (Katz, 1989). In our touch system, the thermal receptors embedded in the skin response with the range of temperature from 5°C to 45°C, and establish the feeling of warmth and cold accordingly (Lederman & Klatzky, 2009). In other words, physical interactions between the skin and touched surface induce the feelings of warm or cold, and the changes of skin temperature at contact is the essential signal for perceiving temperature (Ho & Jones, 2007; see also Ho & Jones's heat transfer model, 2004, 2006, 2008). Within temperature perception, the warm and cold sensations can be considered as distinct sub-modalities, as they are mediated by different afferent fibers (Borhani et al., 2017). Therefore, at least the cold and warm sensations would be distinguished in mental simulation.

In linguistic studies of temperature, it has been found that languages may have a two-term contrast for conceptualization of temperature in linguistic manifestation, which is likely to be cold and warm. It is also possible to categorize temperature in four terms: cold, cool, warm and hot (Koptjevskaja-Tamm, 2015). Although the linguistic variation has been found in temperature, few have investigated the influence of perceiving physical temperature during language comprehension. On the other hand, there are studies assessing the effect of metaphors related to temperature in social experiences. For instance, it seems that our social judgement of other people in term of personality trait "warmth" and the way we treat others were influenced by physical temperature stimulation (Williams & Bargh, 2008). IJzerman and Semin (2009) showed that when participants holding a warm beverage, the social proximity (subjective rate of similarity between the self and others) was higher than holding a cold beverage. Similarly, people felt more "connected" after holding a warm object and warmer after reading positive messages (Inagaki & Eisenberger, 2013), and feelings of belonging were threatened when people with low family support drink cold water (Chen et al., 2014). However, these behavioral studies have received skepticism following failures in replication (e.g., Chabris et al., 2019; Lynott et al., 2017; Wortman, Donnellan, & Lucas, 2014), and hence other paradigms were being sought to investigate the link between social warmth and physical temperature (e.g., Borhani et al., 2017; Fetterman, Wilkowski, & Robinson, 2018). Regardless, these studies did not directly assess whether the processing of language is grounded in the perception of temperature. Instead, they may reflect association between physical temperature and social experiences in daily activities. For example, the temperature-aggression hypothesis suggests people are more aggressive when temperatures increase (Lynott et al., 2017).

Although the sense of touch, including temperature, is the relatively neglected sense in embodied language research, one study has found an implicit thermal-emotional association occurring in semantic and physical thermal representation (Bergman et al., 2014). Still, whether there is a direct link between temperature processing and its corresponding semantic processing has not yet been investigated. To gain more solid evidence on embodied language processing of touch, the study 2 investigated the influence of temperature on both text and speech perception. In the study, the strength of the association between physical temperature and semantic processing was determined. A speeded target discrimination task was conducted. The target stimuli were the words containing the meaning of the thermal qualities of either "warm" or "cold", while the physical thermal stimulation was presented to the participants' hands. The target stimuli divided into two groups (by the meaning of "warmth" or "coldness") were shown in the form of visually or auditorily presented words.

3.1 METHOD

Participants. Ten paid participants (mean age = 32.7, standard deviation [SD] = 9.21; 6 females) participated in the text experiment. Twelve paid participants (mean age = 33.8; SD = 9.06; 8 females) including 10 repeated members took part in the experiment of spoken words and reported the subjective intensity ratings of the target stimuli. All participants had normal or corrected-to-normal vision and none of them had any motor or linguistic abnormalities. The study was approved by the ethics committee of NTT Communication Science Laboratories and conducted in accordance with the ethical standards in the 2013 Declaration of Helsinki. All participants gave their written informed consent before the start of the experiment.

Apparatus. The experimental system (see Figure 3-1) included a set of thermal display, a 19.6-inch liquid crystal display (LCD, L761T-C, EIZO Corporation), and numeric keyboards (NT-19UH2BK, SANWA SUPPLY). The thermal display was composed of two Peltier devices (TEC1-12730, Hebei I.T. Co., Ltd) having a touchable surface size of 62 × 62 mm. There were two electric fans placing underneath the thermal display. The thermal display was controlled by a proportional integral-differential (PID) controller that programmed in MATLAB. The controller monitored the surface temperature of the Peltier devices and maintained the temperature at the desired one. The responses were collected by the built-in functions of Psychtoolbox and two response keys on two numeric keyboards. Participants pressed the response keys using the index finger with the palm placed on the thermal display as shown in Figure 3-2. The distances between the response keys and between the Peltier devices were 13 and 20 cm, respectively. The participants were seated in front of the LCD display with a resolution of 1,280 × 1,024 pixels and a refresh rate of 60Hz. The distance between the participant's eyes and the screen was about 50 cm. The experiment was conducted in a room with air-conditioning to keep the room temperature at 28 ºC.

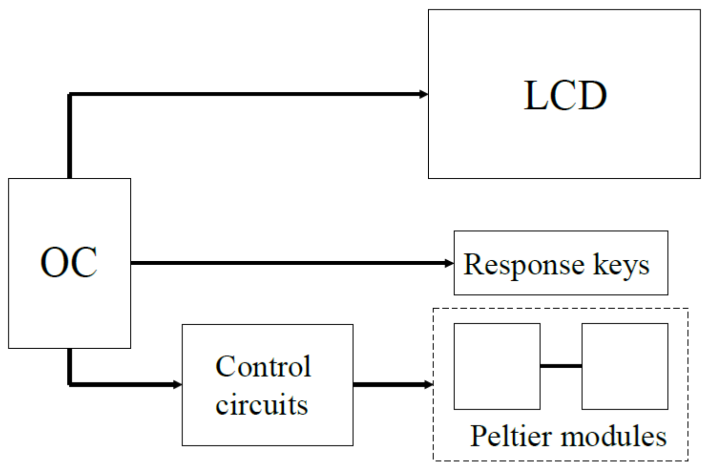


Figure 3-1. Experimental system configurations. OC = Operating computer. LCD = Liquid crystal display.

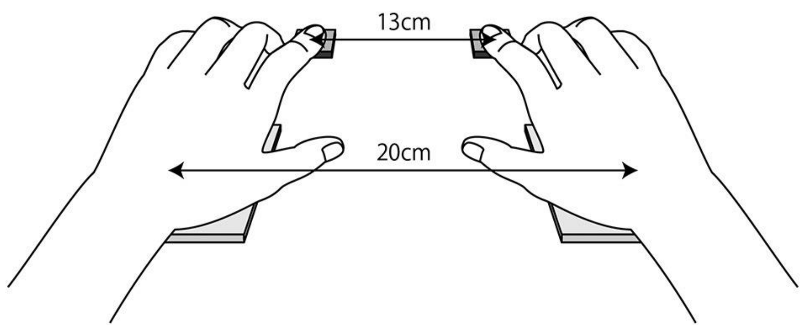


Figure 3-2. Response keys and Peltier devices.

Stimuli. There were 50 words in total and 25 each were selected in two categories: words had the meaning of coldness and the words contain the meaning of warmth (see Table 3-1 and 3-2). The target words used in the experiment were in Japanese because all the participants were native Japanese speakers. The sound recordings of the target words were recorded by a native Japanese speaker in a soundproof room at a sampling frequency of 44.1 kHz.

To confirm the appropriateness of the semantic words in terms of the association with warm/cold concepts, 10 participants who took both the experiment of text and spoken words also evaluated the subjective thermal intensity of each word. The subjective ratings were obtained with a 7-point scale from -3 (extremely warm) to 3 (extremely cold) for "cold" descriptors and from -3 (extremely cold) to 3 (extremely warm) for warm descriptors. The group mean of subjective ratings was 2.24 for "cold" words and 2.49 for "warm" words. These values correspond to the labels between "extremely cold/warm" and "a bit cold/warm" for cold and warm descriptors, respectively, indicating the selection of words shares a common conception of warmth/coldness, at least in the realm of Japanese language.

Three types of physical thermal stimulation generated by the Peltier devices were used: 37 ºC for creating a warm, 27 ºC for cold and 32 ºC for neutral temperature. The temperatures were chosen because they are within the innocuous range of thermal sensation and are similar in terms of subjective thermal intensity (Greenspan et al., 2003). Because the neutral temperature of human skin is 32 ºC, the 5 ºC difference was enough to induce a warm or cold thermal sensation. All participants confirmed that they were able to feel the warm or cold temperature on both palms before the start of experiment.

Table 3-1. List of words associated with meaning of warmth.

No.	Words	Translation
Warmth		
1	火	Fire
2	炎	Flame
3	太陽	Sun
4	夏	Summer
5	火傷	Burn
6	花火	Fireworks
7	日光浴	Sunbathing
8	焼く	Bake
9	マグマ	Magma
10	日焼け	Sunburn
11	熱風	Hot air
12	砂漠	Desert
13	沸点	Boiling point
14	暖かい	Warm
15	オーブン	Oven
16	火祭り	Fire festival
17	熱中症	Heatstroke
18	熱帯	Tropics
19	沖縄	Okinawa
20	ブラジル	Brazil
21	赤道	Equator
22	火柱	Pillar of fire
23	真夏日	Hot summer day
24	発火点	Ignition
25	灼熱	Scorching heat

Table 3-2. List of words associated with meaning of coldness.

No.	Words	Translation
Coldness		
1	水	Water
2	氷	Ice
3	北極	North pole
4	冬	Winter
5	凍傷	Frostbite
6	スケート	Snow skate
7	雪合戦	Snowball fight
8	冷やす	Cool
9	吹雪	Snowstorm
10	しもやけ	Chilblains
11	北風	North wind
12	凍土	Frozen ground
13	氷点	Freezing point
14	寒い	Cold(feeling)
15	冷蔵庫	Refrigerator
16	雪祭り	Snow festival
17	風邪	Cold(illness)
18	寒帯	Frigid zone
19	北海道	Hokkaido
20	ロシア	Russia
21	氷山	Iceberg
22	氷柱	Icicle
23	真冬日	A day temperature below 0°C
24	絶対零度	Absolute zero
25	極寒	Midwinter

Procedure. A speeded target discrimination task for word stimuli presented in the form of visual word or spoken word was conducted. Participants perceived the physical temperature by putting their palms on the Peltier devices' surface. They were asked to memorize the instruction shown on the screen at the beginning of every session, and press the two response keys to discriminate "cold" and "warm" target words as rapidly and accurately as possible once they have seen the words appeared on the LCD display. Before each experimental session, participants confirmed whether they felt "cold" and "warm" on each palm. The distance between a key and Peltier device was adjusted according to the size of participants' hands, and once the set position was fixed, the participants were instructed not to move the palm on the surface of Peltier device during the experiment.

The experiment included twelve sessions that divided into two blocks. There were 50 experimental trials in each session, and these trials began with the presentation of a fixation point in the center of the screen for an interval of 800ms. The target stimulus was then presented centrally until the participant responded. Experimental trials were presented in a different random order for each participant.

Between each session, there was a 3-minute adaptation period that participants were asked to move their palm onto a silicon rubber plate heated to 32 ºC to adjust their skin temperature back to neutral to ensure they would perceive the desired thermal sensation. There was a 30-min break between blocks. The first two sessions in each block were practice sessions, in which the neutral thermal stimulation at 32 ºC was presented to both hands, with the response key for the warm meaning assigned to either the left or right hand and the one for cold meaning assigned to either the right or left hand. The response key assignment was reversed between the two sessions. The third and fourth (or fifth and sixth) sessions were congruent sessions and the fifth and sixth (or third and fourth) sessions were incongruent ones. In the congruent sessions, the warm and cold stimuli were presented to the right and left (or left and right) hands, with the response key for the warm meaning assigned to the hand with the warm stimulation and the one for the cold meaning assigned to the hand with the cold stimulation. In the incongruent sessions, the warm and cold stimuli were presented to the right and left (or left and right) hands, with the response key for the warm meaning assigned to the hand with the cold stimulation and the one for the cold meaning assigned to the hand with the warm stimulation. The location of the temperature stimulation was reversed in two sessions of the same conditions.

To summarize, an experimental block included two practice sessions, two congruent condition sessions, and two incongruent condition sessions. Two experimental blocks were conducted. The order in which a condition (congruent and incongruent) was presented first in the blocks was counterbalanced across participants. Each block took around 30-min to complete for each participant. The RTs and response accuracy of the participants were collected.

The experimental procedure of sound experiment was slightly different compared to the experiment using text as the target. In the sound experiment, there were four blocks and each block contained three sessions, which were practice (i.e., neutral thermal stimulation at 32 ºC was presented), congruent and incongruent sessions. One experimental block took 15-20 minute to finish and the rest time between each block was 20 minutes. All participants performed the 3-minute adaptation period between each session to adjust to a neutral temperature. The subjective intensity rating of words in terms of strength of thermal quality was conducted after participants completed all the sessions of sound experiment.

3.2 RESULTS

The average RTs of the categorization task has been analyzed. Only the results of congruent and incongruent groups were examined. Due to a mechanical failure of the keyboards used in the experimental system, 1.10% invalid data occurred in the semantic words experiment and hence had to be removed. The results of incorrect responses (2.14% of the rest trials) were also excluded. Group means of RT for discriminating target words ("warmth" and "coldness") were represented in Figure 3-3. Using RT of correct responses excluding outliers as the dependent variable, a three-way ANOVA in repeated measures was conducted using the software SPSS Statistics. The within-participants factors of congruency were congruency (congruent/incongruent combinations), meaning of words ("warm" and "cold" words) and the test order of conditions (congruent first/incongruent first). The main effect of congruency was significant, [$F(1, 9) = 18.246$, $p = .002$]; the average response latencies of congruent condition sessions were shorter than incongruent one. The main effect of meaning was not significant, [$F(1, 9) = 4.147$, $p = .072$]; although the average RT of “warm” words was faster than “cold” words. There was no significant effect of test order, [F (1, 9) = .022, p = .886]; in other words, the order in which congruent condition session or incongruent condition session was performed first did not affect the RTs. Consistent with the expectation, no significant interactions were observed as word lengths and word classes were similar in the warm or cold words, [for meaning × congruency, F (1, 9) = 1.497, p = .252]; [for congruency × test order, F (1, 9) = 0.277, p = .611]; [for meaning × test order, F (1, 9) = 0.653, p = .440]; [for meaning × congruency × test order, F (1, 9) = 0.432, p = .528].

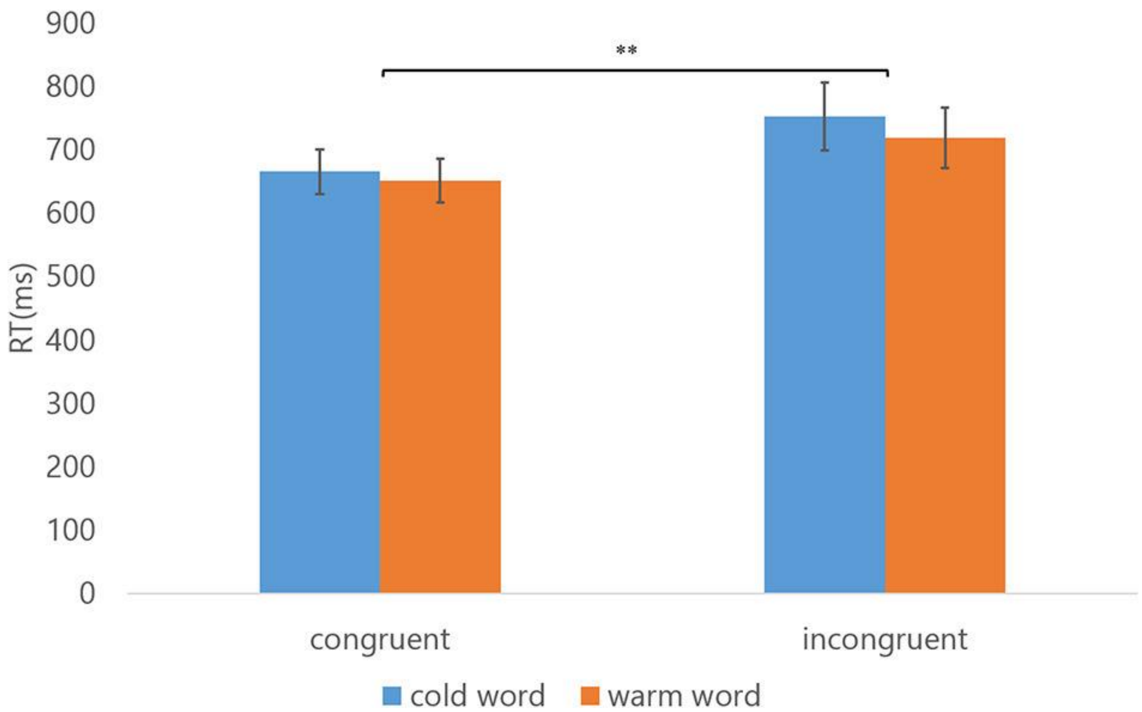


Figure 3-3. Results indicating the mean of RTs (text). The target stimuli in two categories are indicated by the color of the bars, with blue and orange bars representing cold word and warm words, respectively. The error bars show the standard error of the mean. ** indicate statistical significance of $p < 0.01$ (three-way repeated measures ANOVA).

In the sound experiment, the 0.96% invalid data has to be neglected due to the Due to a failure of acquisition from key pressing. After the failure, 1.82% of the data were incorrect in the trials and hence were removed. The group mean of RTs for discriminating the meaning of "warm" and "cold" words presented in audio recordings were shown in Figure 3-4. The within-participants factors of congruency, meaning, and test order were examined by a three-way repeated measures ANOVA using SPSS Statistics software with RT as the dependent variable. There was a significant effect of meaning, [F (1, 11) = 6.761, p = .025]; Participants' RTs were shorter when heard audio recordings of "warm" word than "cold" word. There was no significance found in congruency and test order, [for meaning, F (1, 11) = 3.457, p = .090; for test order, F (1, 11) = .016, p = .903].

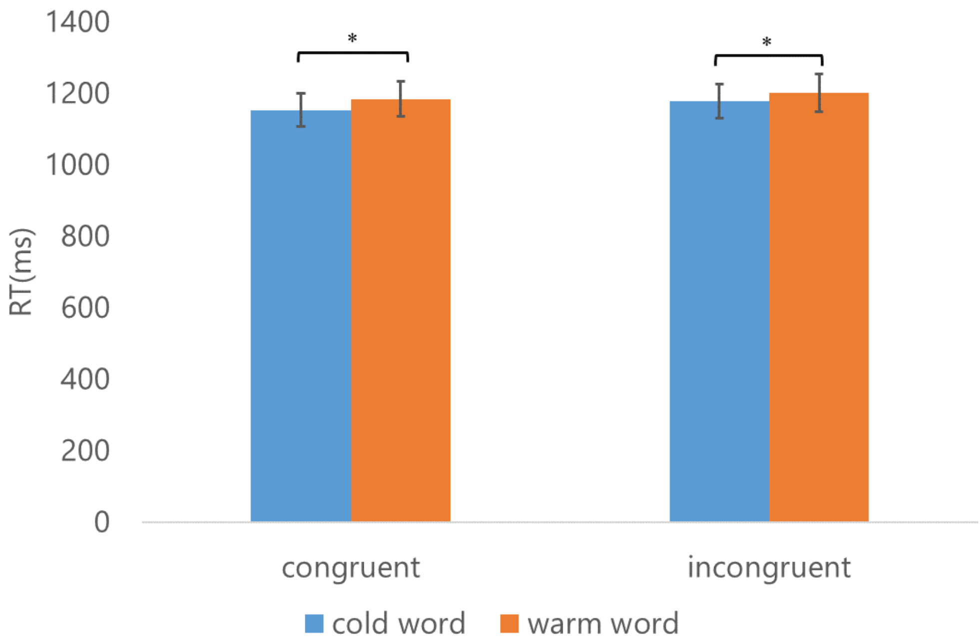


Figure 3-4. Results indicating the mean of RTs (speech). The target stimuli in two categories are indicated by the color of the bars, with blue and orange bars representing cold word and warm words, respectively. The error bars show the standard error of the mean. * indicates statistical significance of $p < 0.05$ (three-way repeated measures ANOVA).

3.3 DISCUSSION

Previous studies have revealed that there may exist a link between physical temperature stimulation and language processing (e.g., Bergman, 2015). Although the relationship between perceptual and conceptual thermal qualities is apparent and intuitive as they display a high degree of similarity, the influence of sensory processing on the internal processing of meaning was undetermined. The aim of the study 2 was to assess the association between physical warmth and coldness and corresponding semantic processing, i.e., whether the physical temperature influences the processing of meaning. The results show a congruency effect on text processing, where the congruent assignment of thermal stimuli to the response key led to shorter RTs than the incongruent key assignment. The perceptual-semantic correspondences influenced the discrimination speed to a semantic-word/illustrative stimulus, and that the "warm temperature-warm meaning" and "cold temperature-cold meaning" congruency facilitated semantic processing. The results are the first to demonstrate an association between physical thermal stimulation and semantic representations in terms of written word using a direct performance task of semantic categorization. In the task, participants were required to press the response key as accurately and rapidly as possible. Therefore, they have to ensure that they correctly understand the meaning of the target words in order to complete the task. All incorrect responses were removed from the analysis. The results directly reflect the influence of temperature in retrieving semantic knowledge, as the RTs in the congruent sessions were facilitated in the experiment of text. Previous research has demonstrated that temperature stimulation cannot influence the lower-level visual processing such as color (Ho et al., 2014), while the results of the present research suggest that complex visual inputs such as text, was influenced by temperature. Therefore, the effect was not a simple cross-modal interaction between visual and tactile modality, and it was more likely to occur at the level of text perception.

The facilitation effect is compatible with embodied language that text processing is grounded in temperature. It has been proposed that the internal representation is implemented via simulations in the brain's modal systems and interfaced with the external structures (Barsalou, 2010). Previous studies suggested that thoughts are implemented in sensory-motor simulations that concepts may be grounded in physical experiences though conceptual metaphor (Lakoff & Johnson, 1980, 1999; Singer et al., 2006; Zhong & Liljenquist, 2006; Jostmann, Lakens, & Schubert, 2009). The present study investigated how external stimulation influences the process of acquiring meaning with semantic representations without any specific conceptual metaphor (e.g., "affection is warmth"), the results indicate that the bodily experiences influenced the processing of semantic meaning presented in text. Moreover, the degree of embodiment might be different between text and speech, as the RTs in the congruent condition sessions were not significantly reduced for spoken word targets. The embodied effect might be observable in processing of text but unobservable in speech at the level of RT measurement.

In addition to the hypothesis of embodied language, it is also possible that attention deployment may play a role in such effect, as the semantic congruency might affect attention control. That is, the internal attention toward the hands located on the thermal display that is congruent with the meaning of the stimulus might be oriented faster than on one that is incongruent with it. Recent research has demonstrated that semantic congruency can affect attention control. For instance, Mastroberardino et al. (2015) showed through fMRI analysis that cross-modal semantic congruency can affect visuo-spatial processing in an audio-visual task. Similarly, behavioral studies have found that cross-modal semantic congruency influenced visuo-spatial attention (Iordanescu et al., 2008, 2010).

The interpretation of interaction between temperature and language processing is supported by the research of the brain's neural processes. There is activity and functional connectivity in the insular cortex when thermal stimuli are processed (Peltz et al., 2011). A meta-analysis (Liakakis et al., 2011) also found effective connectivity between the insular cortex and the area of the inferior frontal gyrus (BA44, the Broca's area for language processing). Not limited to these areas, the insular cortex is broadly connected to other language centers, such as the supramarginal gyrus for language repetition, and may play a role in coordinating higher-order cognitive processing of language (Ardila, Bernal, & Rosselli, 2014; Oh, Duerden & Pang, 2014). Still, more specific investigations of brain connectivity for text processing in language are required.

The present study is a successful demonstration of grounding text processing in perception especially in the temperature. The processing of meaning was facilitated by the experience of physical temperature, as the response latencies of congruent sessions were faster than incongruent ones. Such an enhanced effect may be beneficial in real-life situations. For example, specifically regarding the effects related to temperature, studies have confirmed that the experience of physical warmth influences affective comprehension and behavior. The conceptual thought and language are understood via the conceptual metaphor of "affection is warm/disaffection is cold" (See also Zhong & Leonardelli, 2008; Ijzerman & Semin, 2009; Citron & Goldberg, 2014). The temperature-semantic association demonstrated in this experiment might help us to understand such conceptual metaphor.

CHAPTER 4

STUDY 3: TOUCH INFORMATION (HEARTBEAT VIBRATION/TEARDROP) DID NOT INFLUENCE NARRATIVE PROCESSING

Narrative processing plays an important role in language comprehension, as it is a major medium of daily communication. For example, in everyday life, we read news stories to gain essential information and read fiction for pleasure. In these stories, a sequence of events and a situation are described. A skilled reader would understand the events without apparent effort. The narrative text is not only understood at the level of word or sentence, but also by narrativity (i.e., events or settings described by a connected sequence of sentences). Other than transmitting information or experience in a story, narrative processing is also essential in episodic memory, autobiographical reasoning, and establishing one's own identity (Habermas & Diehl, 2013; Rall & Harris, 2000; Mar, 2011).

During narrative comprehension, readers construct the representations of the character, states, evens, goals and actions that are described by a story. In other words, when comprehending the story, a microworld is created in the brain of the reader and the linguistic structure of this story can be seen as a set of processing cues on constructing such a world. In early research of narrative comprehension, a model that constructing mental representations of text named situation model has been proposed (Bower &

Morrow, 1990; Bransford, Barclay, & Franks, 1972; Glenberg, Meyer, & Lindem, 1987; Johnson-Laird, 1983; van Dijk & Kintsch, 1983). In this model, the comprehension of a story involves not only the mental representation of the text itself, but also the mental representation of the state of affairs implied by the text. The mental representation is conceptualized as a network of representation units that are linked via the arguments that they share.

The later developed event-indexing model specifies the relations among elements of the situation model (Zwaan, Langston, & Graesser, 1995; Zwaan, Magliano, & Graesser, 1995; Zwaan et al., 1998). The basic unit in this model is called event representation. There are five basic dimensions for relating each of the events: time, space, causation, motivation and entity. The event representation is formed when each clause is processed, and it is integrated with the event representation in the working memory based on its overlap on these dimensions. For example, if the event involves the same person or object, there is overlap on the dimension of entity. It is assumed that the effects of basic dimensions are additive, and several empirical studies have found evidence to support this assumption (Therriault, Rinck, & Zwaan, 2006; Radvansky & Zacks, 2014; Zwaan et al., 1998).

From a perspective of embodiment, the events should be grounded in perceptual experiences. To achieve a fluent narrative comprehension, the sensory-motor system is involved to represent perceptual characteristic of an entity when processing a sentence that implies this entity. The strong embodiment theories argue that language processing is directly modulated by sensory-motor system, and the semantic content is "fully simulated" to recreate direct perceptual experiences. In other words, the sensory-motor system is involved in language comprehension at the level of single word, sentence and even narrative processes.

Recently, although several neuroimaging studies showed that sensory-motor systems may also been activated in narrative processing when reading or listening to narrative texts (Chow et al., 2015; Kurby & Zacks, 2013; Nijhof & Willems, 2015), most behavioural findings of embodiment and language over the past 25 years have mainly focused on the processing of individual words or sentences (meta-analysis of these behavioral studies can be found in Louwerse et al., 2015). In general, stimuli used in psycholinguistic studies have typically been decontextualized words and sentences. It leads to the difficulty of extrapolating the patterns of behavioral responses to narrative comprehension because people rarely understand words or sentences in insolation without context. Therefore, it is necessary to take context into account and investigate how embodied representations influence the processing of narrative texts beyond sentential level.

Both of the IEF and indexical hypothesis models are concerned with narrative comprehension (Glenberg & Kaschak, 2003; Zwaan & Ross, 2004). They assumed that perceptual and motor representations are necessary in language comprehension, and the sensory-motor information is recruited in a context dependent way to simulate the situation described in the sentences. For example, the indexical hypothesis model suggests that language refers to the "affordance of a situation (Kaschak & Glenberg, 2000). It also predicts that emotional states influence language comprehension as well as perception and action. According to the framework of indexical hypothesis, the processing of language and emotional system have several pathways to interact. The term "affordances" can be used to explain the influence of emotions in language comprehension. As introduced in Chapter 1, affordances are based on objects and bodily states that explain the interaction between physical and biological system. For example, only humanlike bodies are afforded to sit in an object of chair, and elephants are not. Emotional states may influence language comprehension by changing the bodily dependent affordances, as bodily states are crucially relevant in emotional processing (Damasio et al., 2000; Craig, 2002, 2004). Moreover, the processing of emotional system involves not only the changes in neural systems that contribute to emotion (insula, amygdala, etc.), but also the changes in physiological states (heart rate, hormonal systems, etc.).

There is already certain evidence that language comprehension at sentential level is affected by emotional states. The emotions conveyed by the text (or speech) may match or mismatch the existing emotional states of the reader (or listener), being in a congruent state should help understand the meaning of text (or speech) if experiencing or simulation emotions is a necessary component of language comprehension. It has been demonstrated that language comprehension was facilitated by congruent emotional state when reading sentences (Havas, Glenberg & Rinck, 2007). In the experiment, respond times to sentences describing emotional events when the participant was in a matching emotional state were shorter than mismatching state: pleasant sentences (e.g., "You and your lover embrace after a long separation") were processed faster when smiling; unpleasant sentences (e.g., "The police car rapidly pulls up behind you") were processed faster when frowning. Therefore, there is an association of emotional and sentence processing (see Figure 4-1). However, whether the effect can be found at higher level of comprehension such as narrative processing has not been investigated.

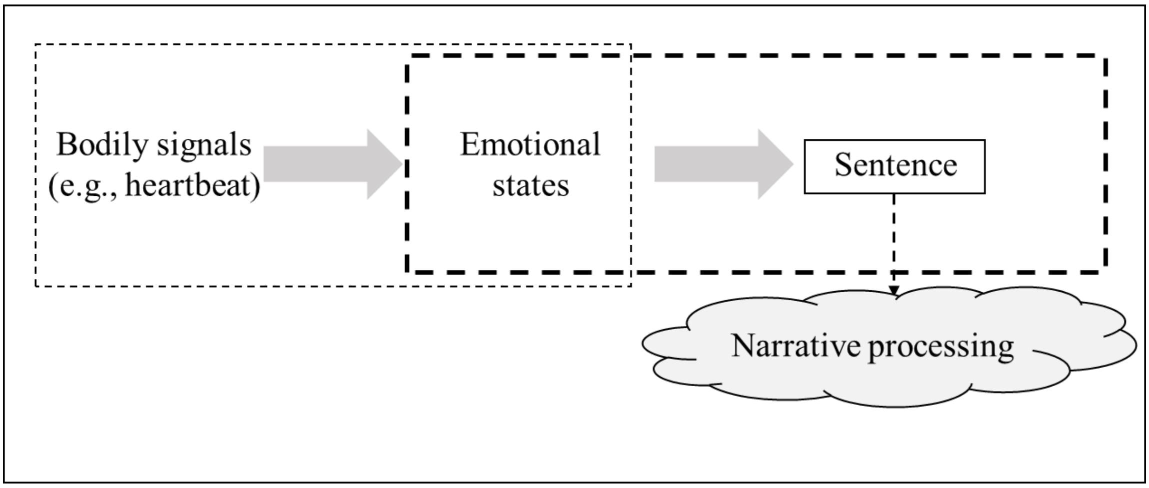


Figure 4-1. Emotional states influence sentence processing.

In addition, among the previous behavioural studies of embodied language, many of them have investigated the contribution of motor actions because actions are central to most human activity. However, other bodily systems such as autonomic, cardiovascular and respiratory systems are also important to cognition and behaviour. For example, the autonomic nervous system (i.e., the system that regulates heart rate, respiration, etc.) contributes to attentional and emotional processing (e.g., Critchley, Eccles, & Garfinkel, 2013; Levenson, 2014). Similarly, the cardiovascular system in charge of distributing blood also contributes to the processing of emotion and wellbeing (e.g., Chapman et al., 2013; Critchley & Garfinkel, 2018). An especially important cardiovascular measure is the heart rate variability because it is associated with this kind of cognitive processing such as emotions (e.g., Ernst, 2017; Kim et al., 2018; Mather & Thayer, 2018; Shaffer & Ginsberg, 2017).

The sensations in the internal body implemented by these bodily systems can be referred as interoception (e.g., heartbeat). The interoception is closely related to emotions. The James-Lange theory is one of the best-known examples of physiological-relatedness to emotion (James, 1884). It argues that emotions are actually induced by the perception of physiological changes, i.e., interoception. Later, other emerging theories of emotions proposed that the emotional responses of a person may be influenced by the specifically induced physiological states. In short, we feel emotions because we perceive our bodily reactions (Schachter & Singer, 1962; Damasio, 1994, 1998; Bennett & Hacker, 2005). There is evidence that the awareness of our bodily signals, such as our heart rate, directly influences the way we feel. For instance, the emotional experience is modulated by perceiving auditory-vibrotactile heartbeat (Tajadura-Jiménez, Väljamäe, & Västfjäll, 2008). There is also neural evidence indicating that emotional experience may be shaped by bodily inputs (Critchley et al., 2004; Damasio et al., 2000).

Although there is a spontaneous relationship between interoceptive information such as heartbeat and emotional/affective states (see Figure 4-2, e.g., we usually experience a rapid heart rate when we are surprised or scared), it remained unclear whether the emotional/affective effect observed in behavioural task is directly induced by the changes in physiological states.

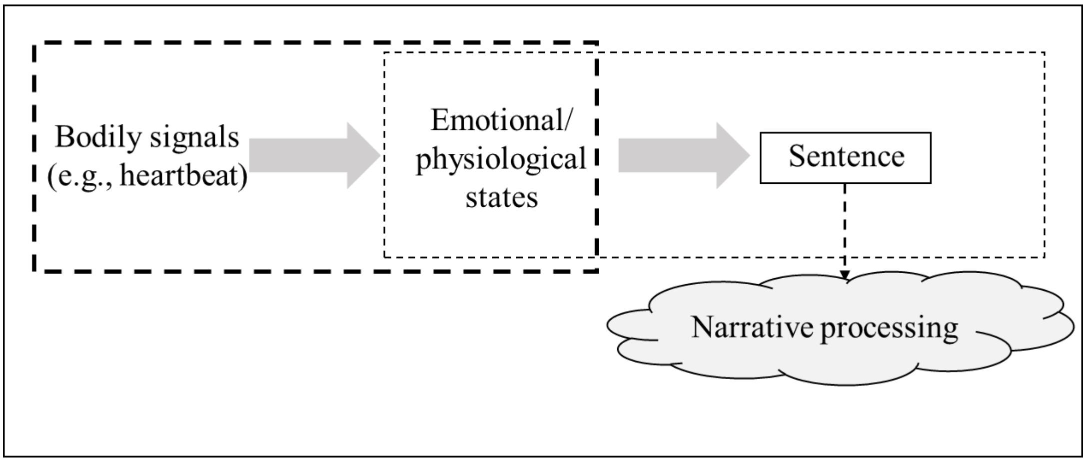


Figure 4-2. Bodily signals (e.g., heartbeat) influence emotional processing.

The role of interception in the emotional concepts has been examined (Connell et al., 2018). Particularly, interoception dominates the concepts of negative emotions such as sadness or fear. Connell et al. (2018) also showed that the strength of interoceptive information contributes to concreteness effect on word recognition (i.e., a semantic facilitation effect that concrete words are processed faster and more accurately than abstract words). As a result, it indicates an implicit access of interoception to the grounding of words. Moreover, a recent study found that primary interoceptive cortex was activated by imagination of being in a situation that related to internal sensations (Wilson-Mendenhall et al., 2019). Although this study only investigated the engagement of interoception in imagination, it indicates the possibility that the interoceptive system is activated when comprehending interceptive language. Therefore, the relation between the interoceptive system and language comprehension should also been examined and included in the empirical assessment of embodied language processing.

In study 3, the effect of perceiving one's own heartbeat in real time on physiological states was first investigated before assessing the influence of such feedback on narrative comprehension of contextual sentences containing affective information (see Figure 4-3). Specifically, whether the heartbeat feedback provided by a haptic device has a positive effect on physiological states was investigated. In other words, whether the heartbeat vibration can induce a relaxation effect was assessed. In addition, the auditory feedback of heartbeat was also examined.

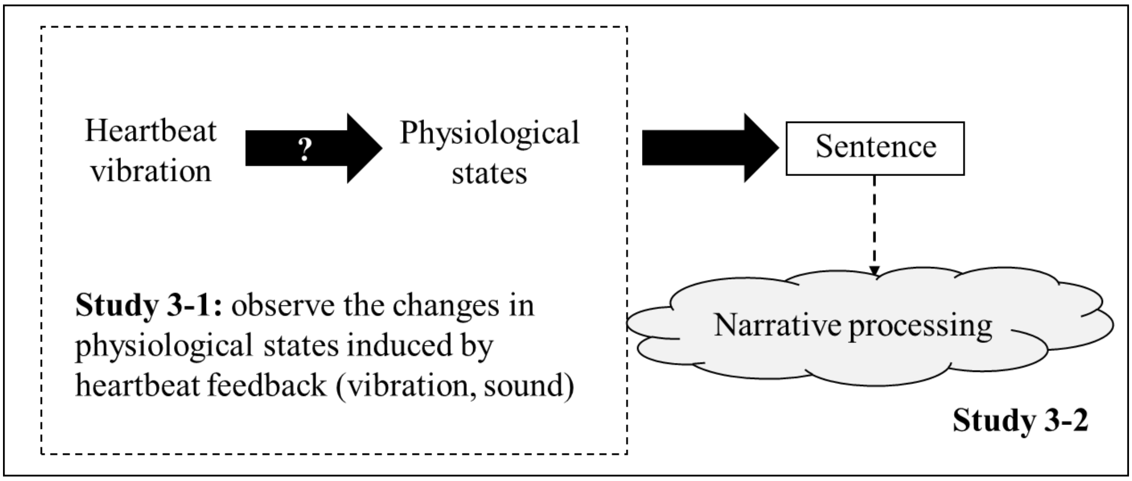


Figure 4-3. The overview of study 3.

4.1 STUDY 3-1: THE CALMING EFFECT OF VIBRATION

Heartbeat, an essential part of cardiovascular system, not only serves as an index of our health status but also plays a crucial role in emotional states. For instance, our heart rate speeds up when we are scared or excited and slows down when we are calm and relaxed. Recent research has applied heart rate feedback to help users regulate emotions. For example, devices such as iAware provide emotional feedback to users according to their heartbeat signal (Albraikan, Hafidh, & Saddik, 2018). Moreover, heartbeat feedback provided via vibration can ease subjective anxiety in stressful situations such as public speech (Costa et al., 2016; Azevedol et al., 2017). Similarly, the calming effect of perceiving one's own heartbeat via vibration has been reported in a workshop called "Heartbeat Picnic" (Watanabe et al., 2011). According to the comments from participants in the workshop, they felt calmer and more focused when they talked to strangers while perceiving their own heartbeats via vibration.

Although it has been demonstrated that heartbeat feedback is efficient in reducing anxiety, i.e., diminishing negative emotions, whether it can foster positive emotions has not been investigated. In contrast to previous studies of easing subjective anxiety, the present study investigated whether the heartbeat vibration induces a positive effect on a relaxing activity (see Figure 4-4). In the relaxing activity, such effect was determined by measuring the changes in physiological states. Physiological states are difficult to manipulate consciously compared with subjective reports of emotional states which may be influenced by individual differences. The investigation on physiological states would provide insights into the involuntary effect of heartbeat feedback on emotional/affective processing. Moreover, considering that auditory feedback has also been frequently used in previous research to investigate the influence of heartbeat feedback on emotional and cognitive processes (Lenggenhager et al., 2013; Tajadura-Jiménez et al., 2008), the effect of auditory heartbeat feedback was also assessed to determine whether it has the similar effect.

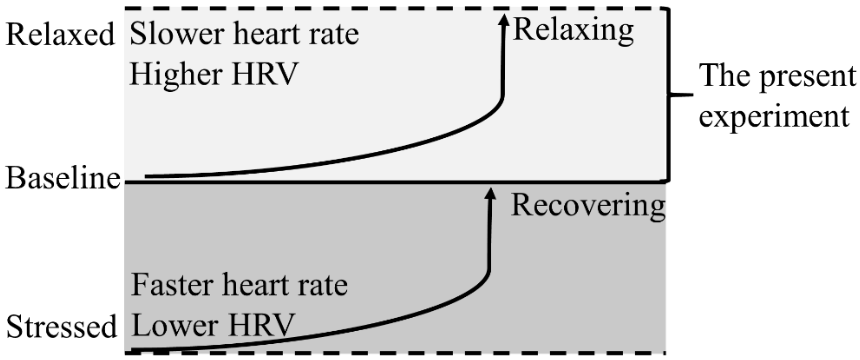


Figure 4-4. The present study investigated the physiological states in a relaxing activity. HRV = heart rate variability. © 2020 IEEE.

The present study used the same device that was demonstrated in the "Heartbeat Picnic" workshop. The device is portable, light, and easy to use. As illustrated in Figure 4-5 (a), the device comprises a stethoscope with a microphone embedded to capture the sound of the user's heartbeat, a vibration speaker (referred to as the heart box), and control circuits to transform heartbeat sound into vibration. To use this device, the user places the stethoscope on the chest with one hand and holds the heart box in the other hand. The user's heartbeat is then simultaneously output as vibration from the heart box. The device can also provide auditory heartbeat feedback by simply feeding the heartbeat signal from the stethoscope to an audio output device. The user can then listen to the heartbeat sound through headphones. Figure 4-5 (c) shows an example waveform of heartbeat signal obtained from the stethoscope.

In this study, the physiological states in terms of heart rate variability (HRV) and heart rate were measured to evaluate the efficacy of heartbeat vibration in emotion induction. In the experiments, electrocardiograms (ECG) signals were recorded while participants were sitting in a chair as shown in Figure 4-5 (b). The HRV is then calculated from the ECG signal to assess the influence of the heartbeat feedback. It is a measure of changes in the time intervals between consecutive heartbeats called interbeat intervals (R-R intervals). According to the Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology (1996), HRV decreases when an imbalance in autonomic nervous system is induced by increased stress. Conversely, increased HRV is usually associated with relaxing and recovering activities. There is a correlation between emotions and HRV (Lane et al., 2009), and HRV may serve as an objective tool to assess emotional responses (Valenza, Lanata, &. Scilingo, 2012).

Two experiments were conducted. Experiment 1 investigated the effect of perceiving participants' own heartbeat via vibration; experiment 2 examined the influence of auditory heartbeat feedback. Three identical sessions were conducted in each experiment to increase the reliability of the physiological measurements. In order to observe the effect of heartbeat feedback, each session included three conditions: first rest, heartbeat feedback (vibration in experiment 1; sound in experiment 2), and second rest. The influence of heartbeat feedback was investigated in a non-emotional situation, i.e., participants perceived the vibration/sound of their own heartbeat while sitting comfortably in a chair without doing any particular task.

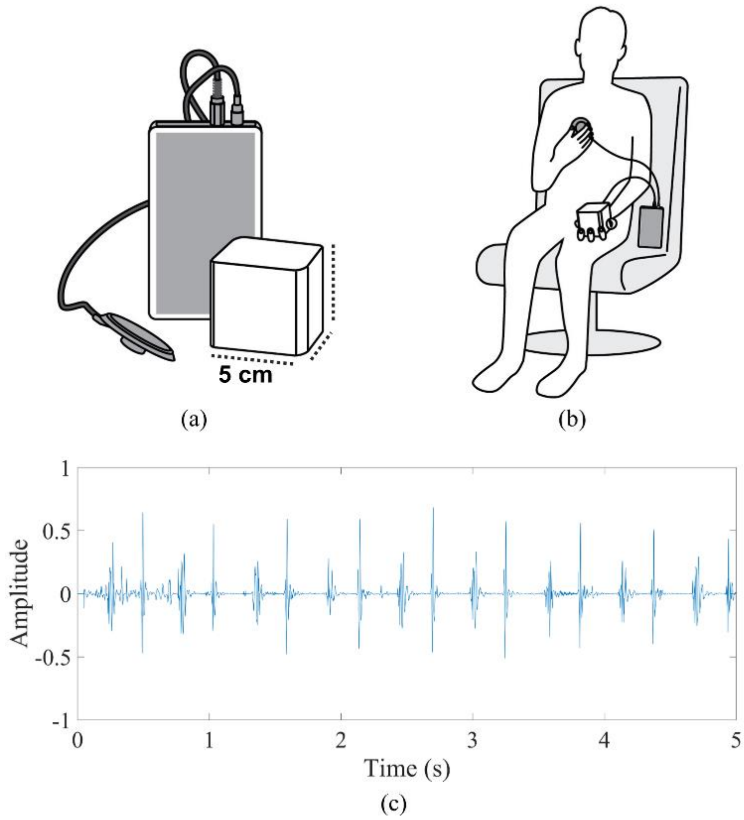


Figure 4-5. The device and the settings. (a) Haptic device for heartbeat feedback. (b) User scene in the experiments. (c) Sample waveform of heartbeat signal obtained from the stethoscope. © 2020 IEEE.

4.1.1 Method

Participants. Twenty-one paid females (mean age = 35.7, SD = 9.3) took part in experiment 1. In experiment 2, 21 people (mean age = 35.1, SD = 9.8), including 19 people from experiment 1, participated. Only female participants were recruited in order to reduce the variation caused by gender (Antelmi et al., 2004). The two experiments were conducted separately on the different days, every participant took at least 2 days (mean = 58.2 days, SD = 63.5 days) break between two experiments. All participants had no known abnormalities of their haptic, auditory, or motor systems and gave written informed consent. The study was approved by the NTT Communication Science Laboratories Research Ethics Committee and was performed in accordance with the ethical standards set down in the 2013 Declaration of Helsinki (Zhou, Murata, &

Watanabe, 2020), © 2020 IEEE.

Apparatus. The haptic device is composed of a stethoscope with a microphone embedded (HBS-NA, YAGAMI Inc.), a heart box with a vibro-transducer (VP408, Acouve Laboratory, Inc.), and our custom-built processing circuits to transform sound signal into vibration signal suitable for the heart box (the sound signals pass through the bandpass filter and are amplified 10 times below 400 Hz, 1 time at 700 Hz and 1/100 times above 2kHz). In experiment 1, the haptic device was used to provide heartbeat vibration. In experiment 2, a handy recorder (ZOOM H2n, Zoom Video Communications) and a headphone (RP-HTX70-H, Panasonic Corporation) were used to provide heartbeat sound. In both experiments, for evaluation of the physiological states, the ECG data measuring R-R intervals were amplified and digitized with a BIOPAC MP150 and two BIOPAC ECG100Cs (BIOPAC System, USA). The sampling rate was 1,000 Hz for all the measurements (Zhou, Murata, & Watanabe, 2020), © 2020 IEEE.

Procedure. Due to space limitation, two participants took part in the experiments in one experimental room at the same time, with a partition placed separating them. Before the experiment, the participants were instructed on how to use the heartbeat feedback device and were given time to experience it. The participants were also asked verbally whether they feel or hear their heartbeat when they experienced the feedback to confirm that the feedback was at a noticeable level. They were then asked to comfortably sit in the chair, keep eyes open and avoid falling asleep during experiment. Both the haptic and auditory feedback experiments included three sessions. The time course of sessions in the vibration experiment is shown in Figure 4-6. In each session, participants were asked to sit still for 10 min (i.e., first rest) to obtain a baseline heart rate, since it has been suggested that participants in an ECG experiment should rest for at least 10 min to stabilize their heart rate before a stimulus is presented (Lee et al., 2010). After 10 min, they were asked to either hold the haptic device or wear headphones to perceive their heartbeats for 5 min (i.e., feedback). Once the heartbeat feedback had finished, participants were asked to remove the devices and complete a questionnaire (i.e., the subjective measurement of anxiety). After finishing the questionnaire, they sat in the chair for another 3 min (i.e., second rest). Participants took a 25-min rest between each session. For the first rest and feedback conditions, the last 3 min of 10- and 5-min recordings were used for the analysis, respectively. For the second rest, whole period (3 min) was used for the analysis. After the experiments, the participants were asked to report whether they fell sleep or felt sleepy during the experiments. Although 9 out of 21 (5 out of 21) participants reported they felt a little bit sleepy at some point during the experiments, none of them reported they actually fell sleep in the experiments. (Zhou, Murata, & Watanabe, 2020), © 2020 IEEE.

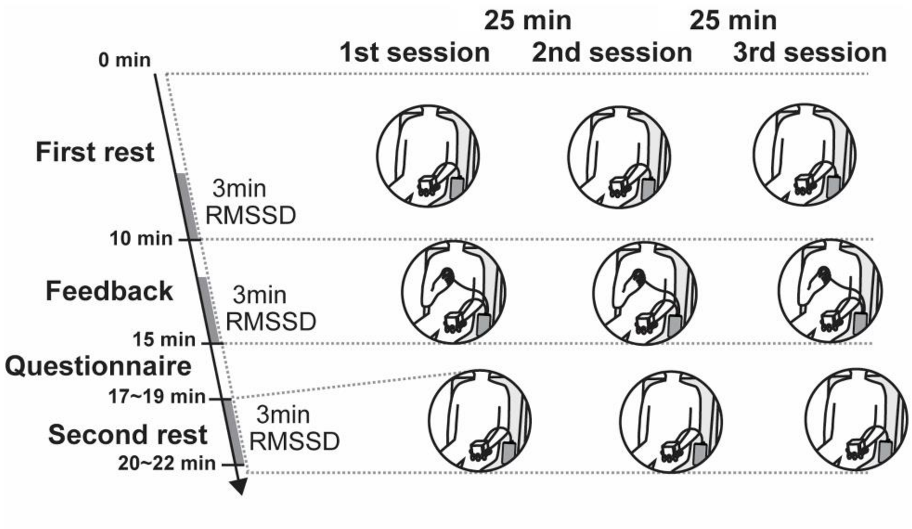


Figure 4-6. Time course of sessions in the vibration experiment. The last 3 min of 10- and 5-min recordings in the first rest and the feedback conditions were analyzed. For the second rest, whole period (3 min) was analyzed. RMSSD = root mean square of the successive differences. © 2020 IEEE.

Data analysis. The two experiments were analyzed separately. In measuring the HRV, one of the most frequently used indices is the root mean square of successive differences between heartbeats, i.e., root mean square of the successive differences (RMSSD). The RMSSD is obtained by first calculating each successive time difference between heartbeats in milliseconds. Then, each of the values is squared and the result is averaged, and finally the square root of the total is obtained. The detailed formula is

$$RMSSD = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n-1}(RR_{i+1} - RR_i)^2} \qquad (1)$$

where RR represents the R-R intervals and n represents the number of RR interval terms.

The RMSSD is a very reliable index reflecting the beat-to-beat variance in heart rate (Shaffer, McCraty, & Zerr, 2014), and it is reliable for various epoch lengths (Wang & Huang, 2012). The higher the RMSSD value, the higher the HRV, meaning that the user is relaxed. In the experiments, a short-term epoch length of 3 min was chosen, since 3-min recordings provide reliable data for most HRV parameters lengths (Wang & Huang, 2012). The R-R intervals were analyzed by AcqKnowledge software (BIOPAC Systems, USA), and the values of RMSSD were calculated from the collected data. Visual inspection and manual correction were carried out to correct misidentified peaks.

For heart rate analysis, the R-R intervals were converted to second-by-second values and expressed as beats per minute (bpm) by dividing 60 by each value. (Zhou, Murata, & Watanabe, 2020), © 2020 IEEE.

Subjective measurement. To measure the influence of heartbeat feedback via vibration/sound on self-reported anxiety levels, participants were asked to complete the STAI-Y-1 questionnaire (Spielberger et al., 1983) before the experiment (i.e., baseline) and immediately after the heartbeat feedback in each session. The STAI-Y-1 questionnaire is composed of 20 items (e.g., "I feel calm"; "I am tense") assessing the state of anxiety. All items are rated on a 4-point scale ranging from "not at all" to "very much so". Scores can range from 20 to 80, and higher scores reflect higher anxiety levels. (Zhou, Murata, & Watanabe, 2020), © 2020 IEEE.

4.1.2 Results

The results of the analyses of the RMSSD and average heart rate in experiment 1 are shown in Figure 4-7 and Table 4-1, respectively. Using the obtained RMSSD as the dependent variable, a two-way repeated measures ANOVA using SPSS Statistics software was conducted. The within-participants factors were conditions (first rest, feedback, second rest) and the order of sessions (first, second, and third). Greenhouse–Geisser correction was used when sphericity was violated. The main effect of conditions was significant [$F(2, 40) = 8.206$, $p = .001$, $\eta 2 = .291$]. A main effect was also observed between sessions [$F(2, 40) = 17.353$, $p = .000$, $\eta 2 = .465$]. No significant interaction was observed between session and condition [F (2.055, 41.093)= 1.145, p= .329, η2= .054]. Bonferroni's method for adjusted pairwise comparison revealed that the RMSSD in the feedback condition was significantly larger than in the first-rest and second-rest conditions [for the first-rest and feedback conditions, p= .009; for the feedback and second-rest conditions, p= .006; for the first-rest and second-rest conditions, p=1.000]. The results also suggest the RMSSD values of third session were significantly larger than that of the second and first sessions, and that the second session had significantly greater RMSSD values than the first one [for the first and second sessions, p= .042; for the first and third sessions, p= .000; for the second and third sessions, p= .008]. The results indicate that participants gradually became more relaxed during the three sessions, and more importantly, that they were relaxed during the vibration feedback condition in each session (Zhou, Murata, & Watanabe, 2020), © 2020 IEEE.

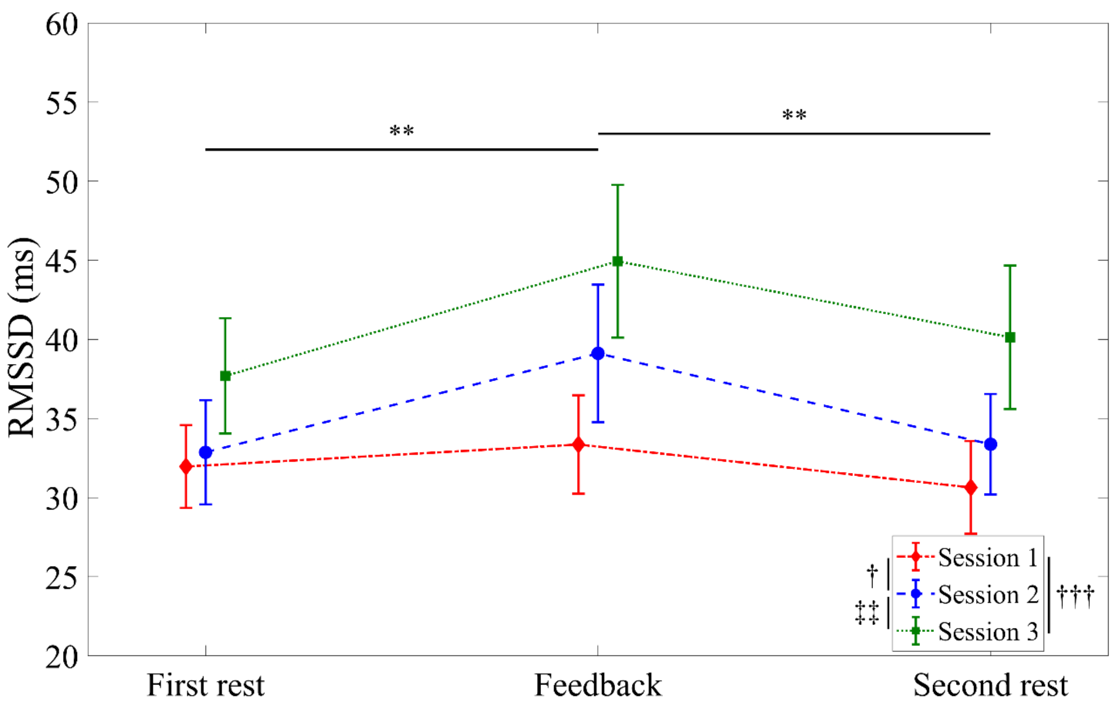


Figure 4-7. Mean values of RMSSD for the three sessions (vibration). Error bars show the standard error of the mean. **p < .01, other conditions versus feedback. †p < .05, †††p < .001, other sessions versus session 1. ‡‡p < .01, other sessions versus session 2. RMSSD = root mean square of the successive differences. © 2020 IEEE.

Table 4-1. Average heart rate during vibration feedback. © 2020 IEEE.

Session	Condition	Bpm (mean ± SE)
1	First rest	70.42 ± 2.74^{f}
	Feedback	69.91 ± 2.90
	Second rest	70.59 ± 2.94^{f}
2	First rest	$68.83 \pm 3.01^{f, s1}$
	Feedback	66.66 ± 2.96^{s1}
	Second rest	$68.36 \pm 2.65^{f, s1}$
3	First rest	$66.56 \pm 2.61^{f, s1, s2}$
	Feedback	$64.78 \pm 2.80^{s1, s2}$
	Second rest	$66.78 \pm 2.77^{f, s1, s2}$

$^{f}p < 0.05$, other conditions versus feedback.

$^{s1}p < 0.05$, other sessions versus session 1.

$^{s2}p < 0.05$, other sessions versus session 2. Bpm = beats per minute.

The average heart rate decreased gradually over the three sessions, as the main effect of session order was significant [F (2, 40) = 14.662, p = .000, η2= .423]. The main effect was also observed between conditions [F (2, 40) = 8.836, p = .001, η2= .306]. No significant interaction was observed between sessions and conditions [F (2.767, 55.349)= 1.229, p= .307, η2= .058]. Bonferroni's method for adjusted pairwise comparison revealed that in vibration conditions the participants had the lowest heart rate [for the first-rest and feedback conditions, p= .012; for the feedback and second-rest conditions, p= .010; for the first-rest and second-rest conditions, p=1.000]. The results also suggest the heart rate values of third session were significantly reduced than that of the second and first sessions, and that the second session had significantly smaller heart rate values than the first one [for the first and second sessions, p= .035; for the first and third sessions, p= .000; for the second and third sessions, p= .047]. These results are consistent with trend that participants became more relaxed during the three sessions, and that, in each session, they became relaxed when they received vibration feedback (Zhou, Murata, & Watanabe, 2020), © 2020 IEEE.

To assess the differences in self-reported anxiety before the experiment (baseline) and after feedback in each session, baseline and feedback STAI-Y-1 scores (see Table 4-2) were entered into a one-way repeated measures ANOVA with timing (pre-experiment, and after feedback in three sessions) as the within-participants factor. Although there was a main effect of timing [$F(3, 60) = 3.581$, $p = .019$, $\eta2 = .152$], the post hoc test with the Bonferroni method revealed that there was no significance between the baseline and after each vibration feedback [for the baseline and feedback in session 1, $p = 1.000$; for the baseline and feedback in session 2, $p = .053$; for the baseline and feedback in session 3, $p = .301$]. Although the p-value for the baseline and session 2 is close to the borderline, no significant effect of the heartbeat feedback via vibration was found on self-reported anxiety in each session (Zhou, Murata, & Watanabe, 2020), © 2020 IEEE.

Table 4-2. STAI-Y-1 scores of the vibration experiment. © 2020 IEEE.

Session	Scores (mean ± SE)
Pre-experiment	35.76 ± 1.41
Session 1	36.05 ± 2.11
Session 2	32.33 ± 1.95
Session 3	32.52 ± 1.97

The RMSSD and group means of heart rate are shown in Figure 4-8 and Table 4-3, respectively. Using the RMSSD as the dependent variable, a two-way repeated measures ANOVA was conducted with the session order and conditions as the within-participants factors. Greenhouse–Geisser correction was used when sphericity was violated. There was no a significant effect of session order [$F(1.404, 28.078) = 4.256$, $p = .036$, $\eta2= .175$] as observed in the vibration experiment. However, there was no significant effect of conditions [$F(2, 40) = 2.704$, $p = .079$, $\eta2= .119$]. No significant interactions were observed [F (4, 80) = 1.495, p = .211, η2= .070]. Bonferroni's method for adjusted pairwise comparison revealed that there was no significant difference between any session pairs [for the first and second sessions, p=.577; for the first and third sessions, p= .101; for the second and third sessions, p=.098]. (Zhou, Murata, & Watanabe, 2020), © 2020 IEEE.

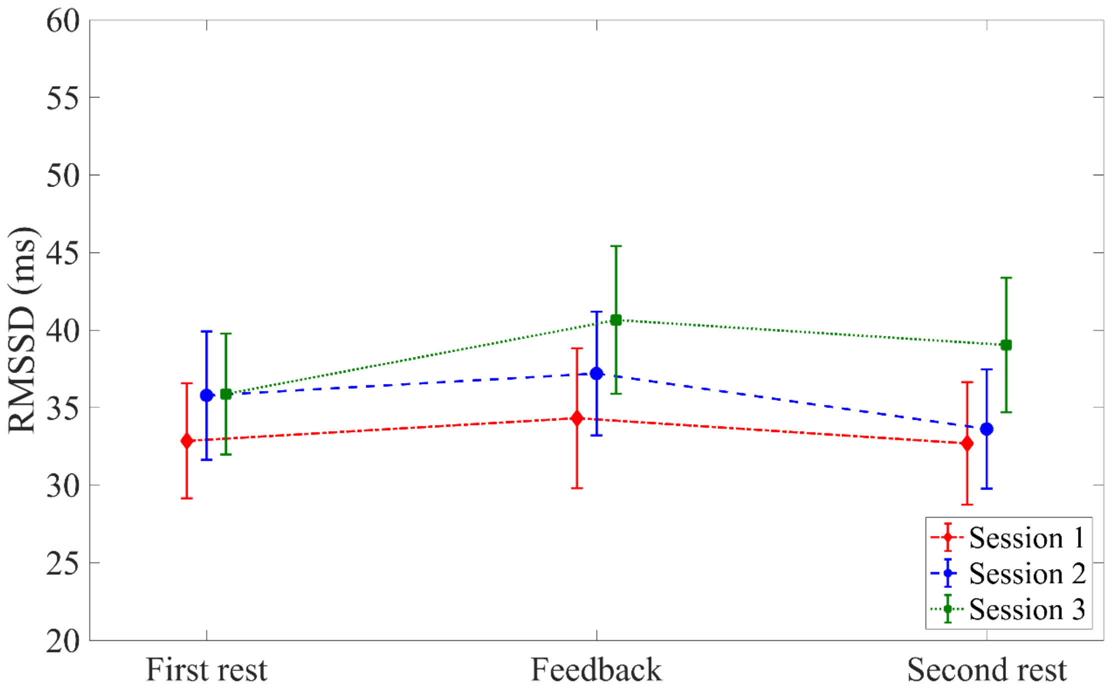


Figure 4-8. Mean values of RMSSD for the three sessions (sound). Error bars show the standard error of the mean. RMSSD = root mean square of the successive differences. © 2020 IEEE.

The analysis of heart rate showed a significant effect of session order and also a significant interaction [for the session order, F (1.432, 28.611) = 13.178, p = .000, η2= .397; for the conditions, F (2, 40) = 2.815, p = .072, η2= .123; for the interaction between session order and conditions, F (4, 80) = 3.040, p = .022, η2= .132]. Bonferroni's method for adjusted pairwise comparison revealed that the heart rate values were significantly reduced in sessions [for the first and second sessions, p= .007; for the first and third sessions, p= .002; for the second and third sessions, p= .046]. The one-way

ANOVA of the results in each session indicate that only in the second session, the second-rest condition was significantly different from feedback condition [p = .005]. (Zhou, Murata, & Watanabe, 2020), © 2020 IEEE.

Table 4-3. Average heart rate during auditory feedback. © 2020 IEEE.

Session	Condition	Bpm (mean ± SE)
1	First rest	71.15 ± 2.85
	Feedback	71.60 ± 2.96
	Second rest	72.16 ± 3.00
2	First rest	69.40 ± 2.71^{s1}
	Feedback	68.66 ± 2.68^{s1}
	Second rest	$70.85 \pm 2.94^{f, s1}$
3	First rest	$68.69 \pm 2.80^{s1, s2}$
	Feedback	$67.40 \pm 2.67^{s1, s2}$
	Second rest	$67.74 \pm 2.82^{s1, s2}$

$^{f}p < 0.05$, other conditions versus feedback.

$^{s1}p < 0.05$, other sessions versus session 1.

$^{s2}p < 0.05$, other sessions versus session 2. Bpm = beats per minute.

Table 4-4. STAI-Y-1 scores of the sound experiment. © 2020 IEEE.

Session	Scores (mean ± SE)
Pre-experiment	32.95 ± 1.71
Session 1	31.48 ± 1.79
Session 2	30.52 ± 1.84
Session 3	30.62 ± 1.86

The same one-way repeated measures ANOVA was conducted to assess the differences in self-reported anxiety before the experiment (baseline) and after feedback in each session. The STAI-Y-1 scores in the sound experiment are shown in Table 4-4. The results indicate that there was no significant difference between the baseline and after each auditory heartbeat feedback, as no main effect was found for auditory feedback [F (1.383, 27.669) = 2.304, p = .133, η2 = .103]. The results indicate that the self-reported anxiety of participants was not influenced by the auditory heartbeat feedback. (Zhou, Murata, & Watanabe, 2020), © 2020 IEEE.

4.1.3 Discussion

The present study investigated the influence of heartbeat feedback of vibration and sound on physiological states and subjective anxiety, In the experiments, no extra emotional load was induced by any additional task, as the participants were only required to sit still in a chair during the experiment. For the results of physiological states, the experiment 1 indicates that there is a calming effect of heartbeat vibration. The values of RMSSD were increased in the feedback condition, which reflect a predominance of the parasympathetic nervous system. Moreover, the average heart rate was also reduced. Therefore, the participants were more relaxed physiologically when perceiving the heartbeat feedback via vibration. In contrast, in the sound experiment, no effect was found in HRV.

In the vibration experiment, there was a 5 to 10 ms difference of RMSSD between the heartbeat feedback condition and the rest condition. This is consistent with previous research using RMSSD as the indicator of physiological states. For instance, in a study investigating the effect of an aerobic exercise intervention among cleaners at work and during leisure (Hallman et al., 2017), the changes in RMSSD were around 3-6 ms. Other studies investigated the influence of meditation on HRV have also reported significant changes in RMSSD about 6 ms (Kulkarni, Nagarathna, & Nagendra, 2010; Steinhubl et al., 2015). Similarly, although the changes in heart rate were small (around 2-4 bpm), they are consistent with other studies. For example, two studies that investigated the influence of music (Ooishi, et al., 2017) and intervention of excises (Hallman et al., 2017) on cardiac activity have reported the significant changes in heart rate around 5 bpm and 2.6 bpm, respectively.

The result of the vibration experiment indicates that external sensory feedback may influence physiological states, specifically the cardiac cycles. This phenomenon has been demonstrated in previous studies, for example, one study found that listening to joyful music evoked an increase in heart rate and a decrease in HRV (Krabs et al., 2015). The phenomenon might be related to the brain's neural processes that feeling emotions requires the participation of brain regions involved in the regulation of internal organism states (Damasio et al., 2000). Particularly, the insula cortex, which plays an essential role in the evaluation and processing of emotions, is activated by cardiovascular arousal (Critchley et al., 2000; 2004). However, the feedback of cardiovascular signal may be processed differently in different modalities, as there was no significant effect found in the sound experiment. (Zhou, Murata, & Watanabe, 2020), © 2020 IEEE.

Despite the significant effect found in the physiological states, the results of subjective measurement indicated that neither haptic nor auditory feedback influenced self-reported anxiety levels. Although previous studies have demonstrated that heartbeat-like vibration reduced self-reported anxiety (Costa et al., 2016; Azevedol et al., 2017), the present study used real heartbeat vibration instead of false feedback received at a slow speed. Moreover, previous research assessed the efficacy of heartbeat vibration on calmness using stressful tasks that induced a high level of anxiety. In contrast, the present study examined the effect of heartbeat feedback in a non-emotional situation (i.e., the relaxation level in a relaxing activity was measured). (Zhou, Murata, & Watanabe, 2020), © 2020 IEEE.

Vibrotactile sensation is one of the most important forms of feedback in the human-computer interfaces used in smartphones, and for gaming and virtual reality. It is also an easy and effective solution for providing biofeedback to the user in real time. Hence, many devices use vibration to provide heartbeat feedback. For instance, there are devices designed using a heartbeat-like vibration to influence the user's emotional states. These devices were developed following the idea that emotions are related to bodily reactions (e.g., changes in heart rate). For example, Nishimura et al. (2012) developed a vibrotactile device that provides a realistic heartbeat vibration to the user. Users of this device believed their heart rate had increased and hence judged semi-nude photographs to be more attractive. In contrast to research investigating the influence of heartbeat feedback on emotional states, other heartbeat feedback devices have been developed specifically for regulating user anxiety. These devices use different approaches to control the user's conscious attention to the intervention. For instance, "doppel", a false heartbeat feedback device designed to reduce the user's subjective anxiety in speech preparation, applies a method to distract the user's attention given to the heartbeat feedback (Azevedol et al., 2017). Users are given a false purpose for using the device and told to neglect the vibration. The subjective anxiety levels are reduced by providing a tailored heartbeat vibration even though the direct attention of participants is not on the vibration. However, "doppel" does not influence the physiological states in terms of average heart rate. In contrast to "doppel", the device used in the present study requires direct attention to the heartbeat feedback. The users were asked to feel their heartbeat vibration. The results indicate that the heartbeat feedback influenced the physiological states but not the subjective measurements of anxiety levels. Therefore, it is possible that the conscious attention to the feedback may have played a role in the calming effect on the physiological and subjective responses. (Zhou, Murata, & Watanabe, 2020), © 2020 IEEE.

Another common approach is to use sound to notify the user of physiological signals. In fact, most of the past studies investigating the influence of heartbeat feedback on emotional and cognitive processes used auditory feedback (Lenggenhager et al., 2013; Tajadura-Jiménez et al., 2008; Valins, 1966). For instance, a study found that heartbeat sounds influence participants' physiological states as well as their emotional judgments of pictures, and recall (Tajadura-Jiménez et al., 2008). In that study, listening to heartbeat sounds increased the participants' heart rate compared with the silence condition, and the emotional responses to affective visual stimuli were also influenced by the heartbeat sounds. However, it must be noted that the false heartbeat sounds were presented at a fixed rate. In contrast to those results, no significant effect was observed in experiment 2 using real heartbeat sounds. These findings suggest that the effect of auditory heartbeat may be related to a specific rhythm or intensity of the feedback. (Zhou, Murata, & Watanabe, 2020), © 2020 IEEE.

In the present study, a calming effect was found on the physiological states using real heartbeat feedback via vibration. This suggests that the heartbeat vibration device could be used to induce a positive effect on physiological states. Clearly, more work using different tasks is needed to determine the exact role that heartbeat vibration plays in moderating physiological states. For example, similar to the anxiety-induced paradigm, it would be interesting to investigate the effect of heartbeat feedback on the recovery after exercises, video games or timed visual search tasks, as these tasks will increase the baseline physiological arousal of the participants. (Zhou, Murata, & Watanabe, 2020), © 2020 IEEE.

To summarize, the present experiments were designed to determine the influence of haptic and auditory heartbeat feedback on physiological states and self-reports of anxiety. The results indicate that participants became more physiologically relaxed when they perceived their heartbeat via vibration, but the effect was not observed using heartbeat sounds. It is necessary to further examine the effects of both real and false vibration feedback on physiological states in different tasks and compare them with auditory feedback. (Zhou, Murata, & Watanabe, 2020), © 2020 IEEE.

4.2 STUDY 3-2: TOUCH INFORMATION (VIBRATION) DID NOT INFLUENCE NARRATIVE PROCESSING

In study 3-1, a relaxation effect on physiological states induced by heartbeat vibration has been demonstrated. The following studies further examined whether such effect can be grounded in comprehension of narrative text. In narrative comprehension (e.g., reading a news story or fiction), a sequence of events and a situation are described. Multiple levels of information are constructed in narrative processing, and it is deeper than understanding representation of word-level or sentence-level. Most of the research on narrative processing has focused on the processing of text, and it has been suggested that readers may construct a situation model during reading (e.g., Kintsch, 1998). This model assumes that a microworld is constructed by the readers when they comprehending a narrative text (i.e., a story) which describes settings and events occurred in the story. In addition to the information that are directly described in the text itself, readers combine information retrieved from prior knowledge in long term memory to make the text complete and interpretable. The dimensions of event representations among elements of a situation model are specified by the event-indexing model (e.g., Zwaan & Radvansky, 1998). According to the embodied view of language processing, the events representations are grounded in perceptual experiences. In addition, when studying embodied language comprehension in context, affect or emotion should also be taken into account (see Barsalou, 2020; Glenberg et al., 2009).

Havas and colleagues (2007) have demonstrated that being in a congruent emotional state facilitates language comprehension of individual sentences. In their experiment, a pen procedure was used to induce typical facial configurations of emotional states. It is a reliable method to induce the state of a particular emotion by manipulating facial configuration that is typical for such emotion. Participants were asked to hold a pen only by using teeth to induce a smile, or only by using lips to induce a frown, and they were naïve about the propose of the holding pen in the experiment. The participants then performed a sentence judgement task. In the task, participants were asked to judge whether the target sentence described a pleasant or unpleasant event by giving the ratings of sentence valence (positive or negative). The results indicate that participants responded to pleasant sentences faster when smiling than frowning, and the opposite was found for unpleasant sentences. In addition, a lexical decision task was conducted using the same settings. However, no effect was found on response times. Therefore, consistent with the indexical hypothesis, the processing of language was affected by emotional states possibly in a way of deriving affordances needed for understanding sentences. That is, when the emotional state is congruent with the content of the sentence, the affordances were derived more easily to comprehend the sentence.

Although it has been shown that emotional states influenced language processing on sentential level, whether such effect can be found on a deeper level, i.e., narrative processing, remained unclear. In addition, it has been confirmed that physiological states were influenced by interceptive information of heartbeat in study 3-1. Therefore, the present study specifically tested whether narrative processing can be grounded in interoceptive information perceived by touch. The target stimuli of narrative text used in the study were the news stories, as the perception of news is a core feature of daily communication and society.

In everyday life, we read news to gain necessary information and keep updated with new events. During the processing of news messages, affect plays a crucial role in information processing (Lang, Newhagen, & Reeves, 1996; Newhagen & Reeves, 1992). Previous research has shown that news stories may elicit affective states such as emotions and moods (e.g., Unz, Schwab, & Winterhoff-Spurk, 2008). According to Russell (2003), mood states are subjectively and temporarily perceived affective states. There are dimensions in mood states, an arousal (from not arousing to highly arousing) and a valence (from negative to positive). The mood states elicited by news may influence the information processing of the news (Nabi & Wirth, 2008). Moreover, the processing of emotional news may influence cognition and memory. For instance, emotional news can be better remembered than neutral news (Miller, 2006). Similarly, news that elicits stronger emotional arousal is learned better than news that does not elicit emotional arousal (Grabe et al., 2003). Therefore, the present study included negative, neutral and positive news stories to investigate the effect of heartbeat vibration on different emotional content. Before the experiment, all target news stimuli were rated by participants to construct a reliable database of emotional news. That is, each news story was verified whether it can be correctly perceived as negative, neutral or positive news. The evaluation of the valence and arousal for emotional news would be reduced if the relaxation effect of heartbeat vibration can be grounded in the news perception.

4.2.1 Method

Participants. Twelve paid participants (10 females) took part in the experiment. The age range of the participants was 20-40 years. All of the participants had no known abnormalities of their haptic or motor systems and gave written informed consent. The study protocol was approved by the NTT Communication Science Laboratories Research

Ethics Committee and was performed in accordance with the ethical standards set down in the 2013 Declaration of Helsinki.

Apparatus. In the experiment, the same heartbeat device as used in study 3-1 was used. The stimulus presentation and response collection were controlled with a personal computer running Psychopy 1.82.01. The responses were made via two response keys on a numeric keyboard (NT-19UH2BK, SANWA SUPPLY). Participants held the device of heartbeat vibration by both hands all the time throughout the experiment except when they pressed the response keys using their right hand. A gentle-on-skin medical tape (Sukinage-t, Nichiban) was used to fixate the stethoscope on the participant's chest.

Stimuli. The database of the emotional news was constructed by creating 80 news stories based on real news, and adapting 40 news stories made by Matsuda, Kaneko, and Taira (2015). The news from Matsuda's work has been rearranged to a format of having 180 words, and separated into 3 complete sentences. A sample of the negative news stimuli can be found in Figure 4-9 (Refer to Appendix A for the 80 news stories created for the experiment). In total, there were 120 pieces of news including 40 pieces of neutral news, 40 pieces of negative news, and 40 pieces of positive news. Before the experiment, another group of nine participants rate the valence and arousal for each news using self-assessment mannequin (SAM) developed by Bradley and Lang (1994). There were nine scales of the ratings from extremely unpleasant (unhappy, angry) to extremely pleasant (happy, delighted) for valence; and from extremely clam (relaxed, sleepy) to extremely excited (activated, vigilant) for arousal.

In preliminary experiment, it has been confirmed that these news stimuli in the database can be correctly perceived as having negative, neutral and positive emotional content. The valence and arousal of all 120 news stimuli were averaged and they were divided into 2 groups with similar ratings (for negative news, group 1, valence: 2.99; arousal: 6.20; group 2, valence: 2.98; arousal: 6.25; for neutral news, group 1, valence: 5.72; arousal: 3.99; group 2, valence: 5.59 arousal: 4.09; for positive news, group 1, valence: 6.82; arousal: 5.48; group 2, valence: 6.82; arousal: 5.49).

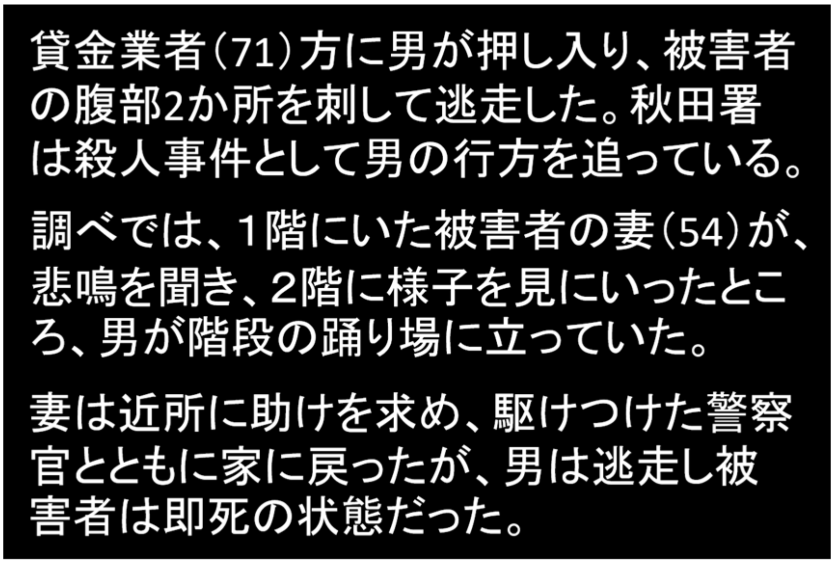


Figure 4-9. News (negative) sample. "A man broke in a money lender' house and stabbed the money lender twice in the belly. Akida police station was searching for the man. According to the investigation of police, the money lender's wife (54) was on the first floor when she heard the scream. She saw suspect standing on the corner of stairs when she was going upstairs, and asked neighbors for help. When she came back to the house with the polices, the suspect has gone and the victim was already dead."

Procedure. In the experiment, participants were asked judge the perceived emotional valence and arousal of the just-read news. There were two conditions, (1) with the heartbeat vibration and (2) without the heartbeat vibration. In the first condition, participant judged the emotional valence and arousal of the news while perceiving their heartbeat vibration in real time. In the second condition, no heartbeat vibration was provided during the news judgment task. In the task, the news was presented in three separated sentences. For each of the news trials, each of the three sentences was presented as text in white for 7,000 ms and then faded out on a black screen. After showing the news stimuli, the SAM appeared on the screen and participants were required to use the keyboard to decide the emotional valence and arousal of the perceived news.

The participants were assigned into two groups of either group A or B. The 120 news trials were divided into two groups. All participants took 120 news trials and came back after 1 week to take part in the exactly the same experiment under the reversed conditions for each news group (see Figure 4-10). The order of the conditions was counterbalanced for each participant group to eliminate the order effect of conditions.

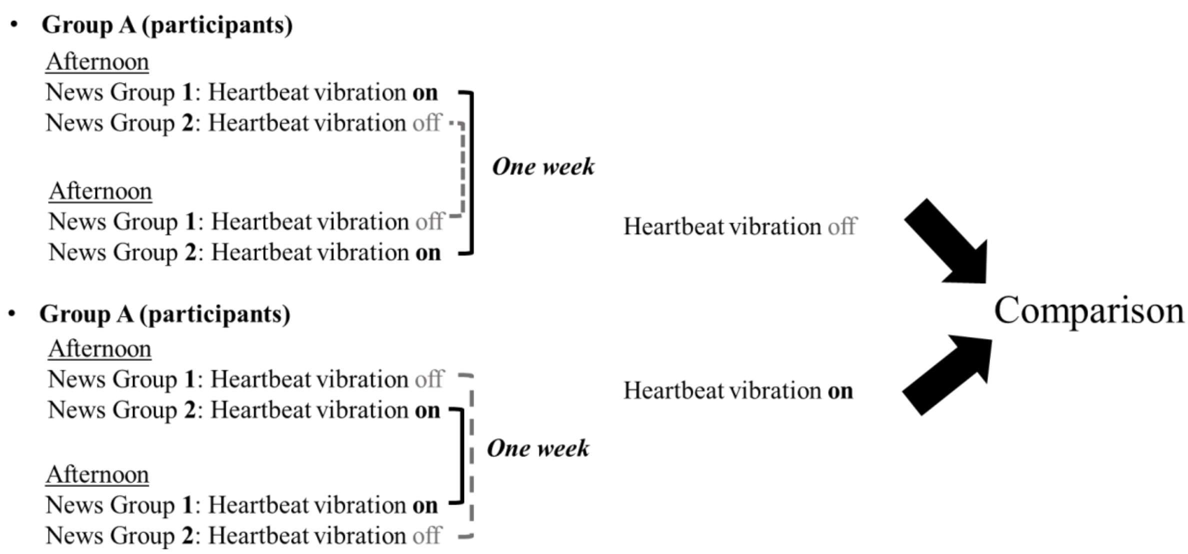


Figure 4-10. Arrangement of the groups of participants and news stimuli.

4.2.2 Results

The group means of the subjective valence and arousal for the two conditions are shown in Figure 4-11. Using the obtained values of valence and arousal as the dependent variables, a two-way ANOVA was conducted by using JASP software. The within participants factors were the conditions (with vibration and without vibration) and the emotional content of news stories (negative, neutral, and positive). For the valence results, no significant effect of the condition was found [F (1, 11) = 1.25, p = .287, $\eta 2$ = .000]. There was a significant difference between news stories in terms of the emotional content [F (1.316, 14.478) = 270.441, p = .001, $\eta 2$ = .936]. No interaction was found between conditions and news groups [F (2, 22) = 0.210, p = .812, $\eta 2$ = .000]; For the results of arousal, no significant effect of the condition was found [F (1, 11) = 0.248, p = .988, $\eta 2$ = .000]. There was a significant difference in the emotional content [F (2, 22) = 48.651, p = .001, $\eta 2$ = 0.275]. No interaction was found between conditions and news groups [F (2, 222) = 0.613, p = .551, $\eta 2$ = .001]. The results indicate that participants correctly perceived the emotional content of the news. However, no influence of heartbeat vibration was observed.

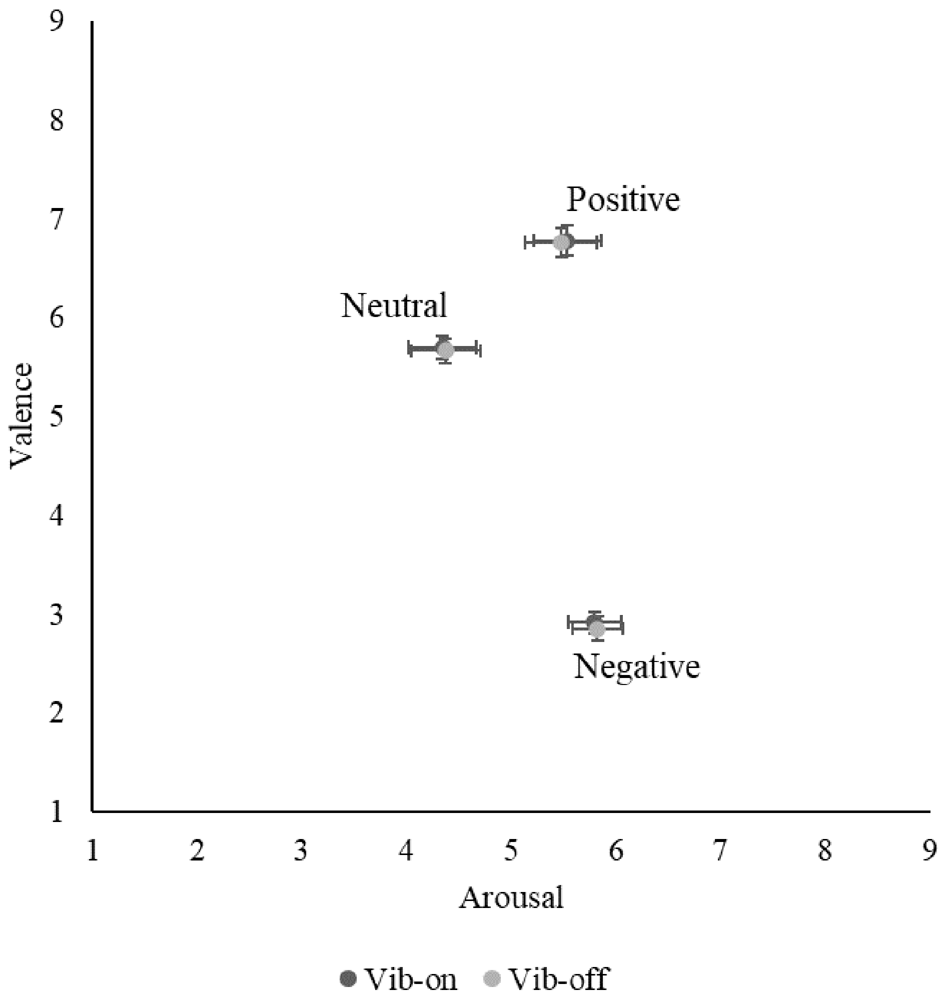


Figure 4-11. Subjective valence and arousal ratings.

4.2.3 Discussion

In the experiment, the influence of interoceptive information, specifically the heartbeat presented via vibration on narrative processing of news stories was assessed. Although there was no significant effect for heartbeat vibration on the judgment of news, a valid database of the emotional news was successfully constructed. Because the heartbeat vibration did not induce a particular emotion, it might be difficult to ground only the changes in physiological states in comprehension of the news stories. Therefore, in the following study, a more explicit stimulus (i.e., teardrop for inducing negative emotions) was used.

4.3 STUDY 3-3: TOUCH INFORMATION (TEARDROP) DID NOT INFLUENCE NARRATIVE PROCESSING

In study 3-3, the influence of touch information of teardrop on the narrative processing was investigated. The same news stories were used to assess the effect of artificial teardrop. The teardrop was provided by a device named teardrop glasses, which was developed by Yoshida (2015). The device is shaped as a pair of glasses, and it can automatically drop tear-like liquid near the lacrimal gland on the face. Negative emotions can be induced by wearing such device. Yoshida (2015) found that participants rated neutral scenes as sad when wearing the teardrop glasses. In the present study, whether the induced negative effect can be grounded in perception of news stories was investigated. The same judgement task of news stories was conducted. In addition, the heart rate variability during the whole experiment was measure to observe whether teardrop induce a change in physiological states. According to the embodiment view of language, the valence and arousal of perceived news were hypothesized to be higher when participants felt the teardrop on the face.

4.3.1 Method

Participants. Four female paid participants took part in the experiment. The age range of the participants was 20-40 years. None of them had no known abnormalities of their haptic or motor systems and gave written informed consent. The experiment was approved by the NTT Communication Science Laboratories Research Ethics Committee and was conducted in accordance with the ethical standards set down in the 2013

Declaration of Helsinki.

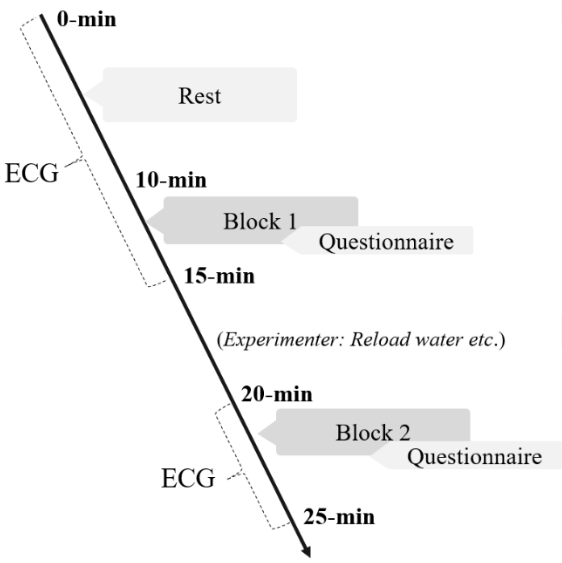
Apparatus. The teardrop glasses used in the experiment were designed and created by Yoshida (2015) using a 3D printer (see Figure 4-12). It had two adjustable pipes to pump the tear-like liquid (heavier than water) on the participant's face. Because of its light weight (72 g), it can be worn comfortably just as the normal glasses. The experience of actual crying was simulated by the tears dropping from the inner corners of the eyes. The maximum duration of the teardrop was about 5 minutes. The speed of the teardrop was controlled by a personal computer (Mac Pro) running Processing, and the interval between two teardrops was at least 2,500 ms.

Figure 4-12. Teardrop glasses created by Yoshida (2015).

Stimuli. The news stimuli used in study 3-2 was selected from the database of emotional news stories used in study 3-1. Only neutral news and negative news were used in the study, and there were 20 pieces for each of them. The perception of teardrop was simulated by the teardrop glasses.

Procedure. In the experiment, participants took the same judgment task as in study 3-2. There were two conditions, (1) judgment with teardrop, and (2) judgment without teardrop. All participants were asked to repeat the experiment by two times, separated by a one-week interval (Day 1 and Day 2). The order in which condition 1 or 2 was first presented were counterbalanced on Day 1 and Day 2 across participants. Four experimental blocks were separated into 2 sessions on each day (see Figure 4-13). Each block included 10 trials of the news stories, and it lasted for about 5 minutes. After finishing one block, participants completed a simple questionnaire to indicate their current mood (i.e., by answering "how do you feel at the moment?" from 1 = very bad to 9 = very good).

At the beginning of each session, predicants rested for 10 minutes to stabilize their heart rate for reaching a baseline status. The measurement of ECG started once the experiment began. Participants took a 25-min rest between each session.



Day 1				
Session	**Block**	**News stimuli**	**Condition**	**Trial**
1	1	Neutral Group 1	No tear	10
	2	Neutral Group 2	No tear	10
2	3	Negative Group 1	Tear	10
	4	Negative Group 2	Tear	10
Day 2				
3	5	Neutral Group 1	Tear	10
	6	Neutral Group 2	Tear	10
4	7	Negative Group 1	No tear	10
	8	Negative Group 2	No tear	10

Figure 4-13. The procedure of the experiment in study 3-3.

Data analysis. Both physiological and subjective measurements were analyzed. For physiological measurement, the HRV and heart rate were calculated. The judgment of subjective valence and arousal for each news story and the scores obtained in the mood questionnaires were analyzed for subjective measurement.

4.3.2 Results

The group means of the subjective valence and arousal for the two conditions are shown in Figure 4-14. Using the obtained values of valence and arousal as the dependent variables, a two-way ANOVA was performed by using JASP software. The within participants factors were the conditions (with teardrop and without teardrop) and the emotional content of news stories (negative, neutral). For the results of valence, no significant effect of the condition was found [$F(1, 3) = .055$, $p = .830$, $\eta^2 = .000$]. There was a significant difference between news stories in terms of the emotional content [$F(1, 3) = 79.731$, $p = .003$, $\eta^2 = .929$]; No interaction was found between conditions and news groups [$F(1, 3) = .010$, $p = .925$, $\eta^2 = .000$]; Similarly, no significant effect of the condition was found for arousal results [$F(1, 3) = 3.125$, $p = .175$, $\eta^2 = .008$]. There was a significant difference in the emotional content [$F(1, 3) = 7.546$, $p = .071$, $\eta^2 = .186$]. No interaction was found between conditions and news groups [$F(1, 3) = .400$, $p = .572$, $\eta^2 = .003$]. The results indicate that participants correctly perceived the emotional content of the news. However, no influence of teardrop was observed.

For mood rating, no significant effect of the condition was found [$F(1, 3) = 0.231$, $p = .664$, $\eta^2 = .001$]. The RMSSD was also not significant [$F(1, 3) = 0.425$, $p = .561$, $\eta^2 = .011$].

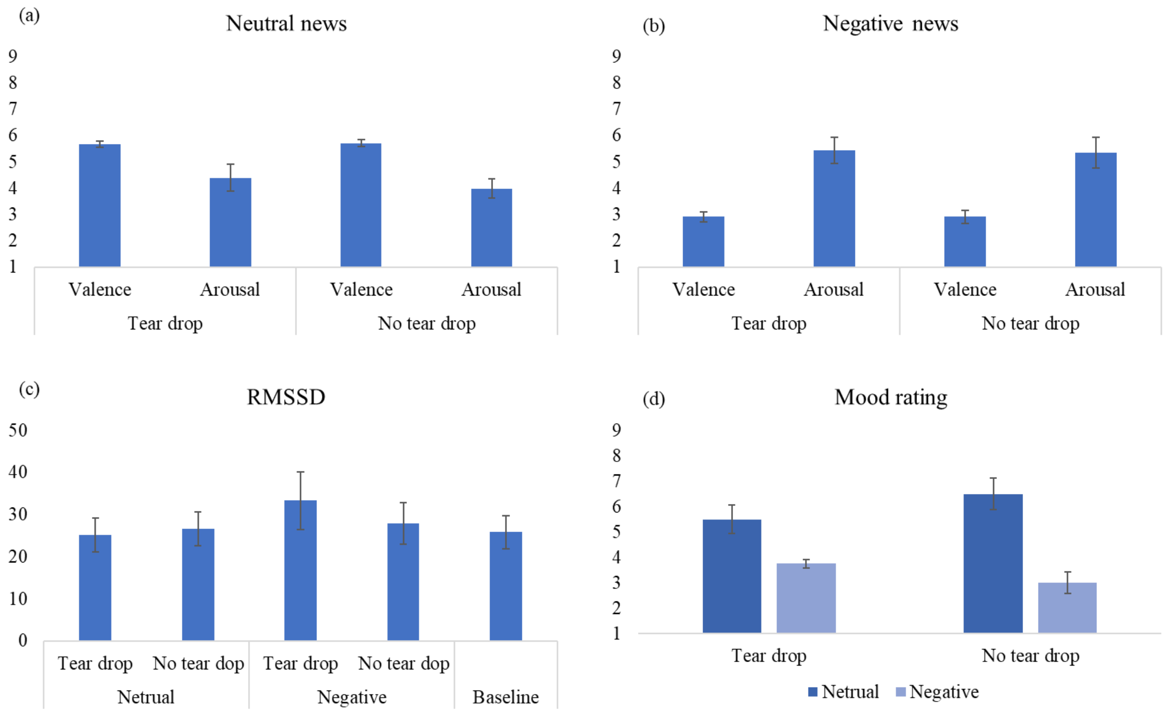


Figure 4-14. Results of study 3-3. (a) Neutral news. (b) Negative news. (c) RMSSD. (d) Mood rating. RMSSD = root mean square of the successive differences. The error bars show the standard error of the mean.

4.3.3 Discussion

Similar to the results in study 3-2, no significant effect of teardrop was found. In contrast to the experiment of Havas and colleagues (2007) that asked participants to grip a pencil to produce a “own” smile on their faces, participants in the present experiment do not physically interact with the teardrop glasses by wearing them (no facial muscle movement involved). Therefore, it might be difficult to induce a sense of ownership for the tears on the face. Previous study has revealed the importance of body ownership in facial feedback phenomena that participants interpreted the robot’s facial expression as their own when they felt ownership of the robot (Nishio et al., 2018). Therefore, having ownership of the facial expression may be essential for producing an embodied effect on cognition.

4.4 DISCUSSION

The results of the study 3 indicate vibration influences physiological states. The sense of the internal physiological states plays an important role in many theories of emotions such as the famous James-Lange theory. According to the James-Lange theory, emotion is experienced as a result of physiological arousal in the body (James, 1884). Therefore, the changes in physiological signal such as heart rate, respiration and skin conductance may induce an emotional experience. This study particularly focused on the influence of heartbeat perception on emotional processing, as the heartbeat feedback can be easily perceived through a haptic device.

Previous studies have already demonstrated a close relationship between heartbeat perception and emotional experiences. For instance, Valins (1966) found that the auditory feedback of false heartbeat influenced the subjective attractiveness of pictures of semi-nude women. Subjects in both increased and decreased auditory false heartbeat feedback groups showed a significant difference in terms of perceived attractiveness compared with the control group. Consistent with the previous works, the results of the present study indicate that interoceptive information, specifically the heartbeat feedback perceived via touch, influenced the participants' physiological states. However, the evidence that perceptual experiences influence narrative processing is insufficient. This suggests it might be difficult to ground perceptual experiences in narrative processing of contextualized sentences. The embodied effect might not be robust beyond the sentential level.

CHAPTER 5

CONCLUSIONS

5.1 SUMMARY AND CONTRIBUTIONS

The aim of this thesis was to investigate to what extent and how language comprehension interacts with sensory-motor systems by examining the influence of perceptual experiences of vision and touch on the perception of speech and text. For speech perception, the involvement of visual perception at the level of component parts of a word/single word was investigated. In contrast, the influence of touch on sentential and narrative level processes was assessed for the perception of text. Three studies are included in the thesis.

The study 1 assessed whether the influence of visual features on spoken word processing can be observed at the sublexical level, the results confirmed that visual information interacts with the perception of auditorily presented geminates. In the experiment, a minimal pair of Japanese words containing singleton/geminate consonants that have the same phonemes but different closure durations, were used as target stimuli. The closure duration of a geminate word is significantly longer than the singleton word. Therefore, the longer the closure duration, the more likely the target is perceived as a geminate word. The participants performed a categorization task for a series of target spoken words preceded by auditory or visual primes. The perceptual bias determined by the shift in perceptual boundary between singleton and geminate (i.e., how rapidly the perception changes) was observed for both prior auditory and visual (written words and illustrations) information. Significantly, consistent with the view of embodied language, the illustrations implying visual features of the objects influenced the spoken word processing at the level of consonants in a similar way as the identical written words do.

Touch, the comparatively neglected sense in empirical research of embodied language, was investigated in study 2. It has been demonstrated that there is an implicit association between emotional concepts and temperature processing. The study 2 further examined whether there is a direct link between the perception of physical temperature and the processing of text. In other words, the experiment in study 2 investigated whether temperature stimulation influences processing of words. In this experiment, participants performed a speeded target categorization task for semantic words representing the thermal qualities of "warm" or "cold", while perceiving the warm and cold temperature. The average RTs for the congruent and incongruent sessions managed by the response key assignments were compared. In the congruent condition sessions, the response key for the words associated with the meaning of warmth or coldness was assigned to the hand attached with warm or cold temperature, respectively. And in the incongruent condition sessions, the key assignment was reversed. The results showed that the response times of congruent sessions were shorter than incongruent ones in written word trials but not in spoken word trials. Therefore, the perception of temperature influenced semantic processing of text, as experiencing the physical temperature facilitates the processing of meaning of thermal quality.

In addition, the study 3 examined the influence of perceptual experience on narrative processing using another "sub-sense" of touch, vibration. In study 3, the influence of vibration coupled with another type of bodily sense, i.e., interoception, was investigated. It has been demonstrated that emotional concepts may be grounded in interoception, although it is not technically a component of touch. The results of the study 3 indicate that interoceptive information of heartbeat presented via vibration induced a relaxation effect on physiological states. However, there was no embodied effect of interoceptive information on narrative processing. Similarly, no effect was found on the processing of the same narrative news when perceiving teardrop on the face. The results of the study 3 indicate vibration influences physiological states, but the evidence that tactile information influences narrative processing is insufficient. Thus, it might be difficult to ground perceptual experiences in narrative processing. The task requirements may also influence the results as the subjective feelings of participants were evaluated instead of measuring the RTs.

The results of the present research are in line with the embodied theory of language comprehension. In particular, the results demonstrated that the influence of perceptual experiences on language comprehension may be different between text and speech. There was no effect of temperature on identical spoken words in study 2. Moreover, both heartbeat vibration and teardrop did not influence the narrative processing. These results reflect a partial dependence between language comprehension and sensory-motor system, and it seems unlikely that perceptual and conceptual processes are carried out by the same mechanism. Therefore, it is plausible to reject the "fully symbolic" amodal view of language processing, although a "full" simulation of perceptual experiences remains doubtful. Among theories of embodied language, the present results are incompatible with the theories classed as strong embodiment that directly predict an influence of sensory-motor systems on narrative comprehension. Instead, weakly embodied theories are supported by the present results. In weakly embodied theories, the language comprehension is partly dependent on perceptual and motor activity.

To summarize, the present research further investigated the influence of perceptual experiences on language comprehension. First, most previous research demonstrating the influence of perceptual experiences on language processing has focused on individual words and sentences. In contrast, the current study showed that the influence of visual information on spoken word processing can be observed at a sublexical level. Second, although touch is a very important sense, relatively little was known about the effect of touch on language processing. The present research investigated the relation between touch and semantic word processing, and has successfully demonstrated a direct link between the perception of temperature and the comprehension of text. Finally, although the evidence of influence of tactile information on narrative processing is insufficient, the present study has shown that heartbeat feedback perceived via vibration influenced physiological states and induced a relaxation effect. In addition, a database of emotional news stories has been constructed. Consistent with the weak embodiment view of language comprehension, the results of the present research suggest that language processing is at least partly grounded in sensory-motor systems.

5.2 FUTURE WORK

Alternative to the embodied language hypothesis, there are other possible accounts for the effects observed in the present study. In study 1, the possibility that the priming effect found in illustration is induced by the shared semantic information between the primes and targets cannot be ruled out completely. A strong priming effect has been demonstrated in the recognition of auditory targets (e.g., salt) preceded by semantically-related visual primes (e.g., pepper) (Anderson & Holcomb, 1995). It indicates that semantically-related words can prime target words across modalities, although in the present study the primes and targets are more than semantically "related": they contain the identical meaning. In addition, speech rate or length of preceding vowel may affect the perception of geminates. The robustness of the priming effect may be assessed by placing carrier sentences before and after the target stimuli to regulate the speech rate within the participants.

In study 2, the bidirectional relation between physical temperature and semantic processing should be verified. Although one study has shown that semantic congruency facilitated haptic qualities of weight, firmness, and vibration (Hecht & Reiner, 2010), the bidirectional relation between physical touch and semantic processing remains to be determined. The interaction between temperature and text processing may also be interpreted as the effect of the attention deployment. That is, the possibility that the semantic congruency influenced attention control cannot be ruled out.

Moreover, there was no facilitation effect on spoken words in study 2. Firstly, the embodied effect was not a simple cross-modal interaction between visual and tactile modality. Previous research has shown that visual processing at the lower level such as the perception of color, was not influenced by the physical temperature stimulation (Ho et al., 2014). Secondly, it is known that the cognitive processes differ in text/speech perception. Therefore, there are two possible interpretations for the disappearance of the facilitation effect in spoken word targets. The degree of embodiment might be different between text and spoken word processing, which was either observable or unobservable in RT measurement. Alternatively, meaning of some spoken words may be easier to acquire: participants pressed response key once they understood the meaning of words.

Therefore, future research might investigate the how speech and text processing differ in temperature-semantic congruency and perform direct statistical comparison between them.

Clearly, more work is required to determine the influence of perceptual experiences on narrative processing. Although the results of the study 3-1 have shown that interceptive information perceived through touch influenced physiological states, it seems that it was difficult to ground the changes in physiological states on narrative processing. The present research applied an indirect stimulus coupling vibration with interoceptive information. The vibro-tactile stimuli were chosen because previous research has demonstrated that emotional concepts may be grounded by interoception; and heartbeat vibration may influence emotions and feelings. The dominate senses such as vision might be used to determine the exact role of direct perceptual experiences in narrative processing.

Finally, the present research focuses on native language (L1) processing of Japanese. In the experiments, all participants were native Japanese speakers and the target stimuli used in the experiments were presented in Japanese. Some bilingualism models assume semantic representations are shared between L1 and L2 (e.g., Bilingual Interactive Activation Plus model [BIA+], Dijkstra & van Heuven, 2002). As a result, L2 is assumed to be affected by perceptual experiences. Although only a few studies investigated the influence of perceptual experiences on L2 processing, both behavioral and neurophysiological studies have found evidence that L2 is very likely embodied (Bergen et al., 2010; Buccino et al., 2017; Xue et al., 2015). The embodied effects might be different in L1 and second language (L2) processing. Some studies have statistically compared the difference in embodied language between L1 and L2 processing, yet, no clear picture has emerged. For instance, Dudschig et al., 2014 found no difference between the embodied effect of L1 (German) and L2 (English) processing in a behavioral experiment, while Qian (2016) demonstrated that the effect was stronger in L1 (Chinese) than in L2 (English). Although there is no firm conclusion of the degree of L2 embodiment, L2 processing is very likely to be grounded in perceptual experiences at least to some extent. The set of the present result might be used to descriptively/statistically compare the difference in embodiment between L1 (i.e., Japanese) and L2 processing (e.g., Chinese or English).

APPENDIX A: NEWS STIMULI

1. NEWS STIMULI (NEGATIVE)

群馬県渋川市の家の庭に生えていた有毒植物「イヌサフラン」を食べて入院、意識不明となっていた男性が死亡した。

「イヌサフラン」は、葉が食用の「ギョウジャニンニク」とよく似ており、過去にも死亡事故があった。

警察は男性が「イヌサフラン」を食用の野草と間違えて食べたとみて、遺体の司法解剖を行うなど死因の特定を急いでいる。

福井県の国道158号で、軽乗用車を積載して走っていた、自動車解体業男性（71）の車両運搬車が横転した。

軽乗用車内にいた、無職女性（66）が胸などを強く打ち、約2時間後に搬送先の病院で死亡が確認され、

同じ軽乗用車に乗っていた女性の夫（69）が右足の骨を折るなど重傷。

浜松市浜北区大平の新東名高速道の橋脚付近で「クマを見た」という情報があり、地元住民らが警戒している。

住民らによると同高速道浜松サービスエリア上りから西に約1キロの雑木林で子グマらしき動物2頭が歩いている。

住民の相談を受けた自治会役員が隣保班長を通じて地区内の100軒以上に情報を伝え、同市や浜北署に通報した。

東京都檜原村の山中にある林道脇の崖下に「軽自動車が転落している」と通行人から１１９番があった。

警視庁五日市署によると、車内にいた無職、同村在住のの82歳と86歳の男性２人の死亡が現場で確認された。

同乗していた７０代と８０代の妻２人も病院に搬送されたが、命に別条はないという。

神奈川県平塚市の警報機、遮断機付きの踏切で、熱海発の回送電車と乗用車が衝突した。

乗用車の大磯町の男性（３１）が病院に搬送されたが、間もなく死亡した。平塚署が事故原因を調べている。

署によると、電車が踏切を通過する直前に乗用車が北側から南側に向けて踏切内に進入したとみられる。

千葉県の東武野田線柏駅で25日午前7時15分ごろ停電が発生し、通勤客ら約5万人に影響が出た。

通勤・通学時間帯を直撃した停電に、東武野田線の柏駅では入場規制が行われ、

改札の前や真っ暗なホームには電車の再開を待つ人や、があふれていた。混乱を防ぐため警察も出動する事態となった。

東京・文京区のビルで、料理教室中にベビーカーが燃え、3歳の女の子が死亡した。

警視庁によると、料理教室の最中に、参加者の母親が室内に持ち込んだベビーカーが燃え出したという。

ベビーカーに乗っていた女児（3）が顔や腕にやけどを負い、病院に搬送されたが、翌日、死亡が確認された。

約７千万円を脱税したとして、元日本弁理士会副会長の弁理士が、東京国税局から所得税法違反罪で告発された。

この弁理士は、知人に設立させた香港のペーパーカンパニーに翻訳業務を委託したように装い外注工賃を計上した。

同社名義の香港の預金口座に振込み、国内のATMから現金を引出す手口で約１億５千万円の所得を隠した疑い。

三重県四日市市のバス会社で、二種免許を持たない管理職の男性社員（55）が貸し切りバスに客を乗せ運行していた。

同バス会社が契約する企業の従業員を貸し切りバスに乗せ、近鉄四日市駅から市内の工場まで2度にわたり運行した。

朝、運転手が手配できていないことがわかったため、大型一種免許しか持たない社員が、運行したということだった。

岐阜市内のアパートの１室で、県警の警察官が男女２人の遺体を発見した。身元と死因を詳しく調べている。

遺体は死後数カ月経っているとみられ、目立った外傷や、部屋が荒らされた様子はないという。

岐阜市の職員が訪問したところ連絡が取れず、署に通報したという。

茅ヶ崎市で、パトカーに追跡されていた少年(18)が運転する2人乗りのバイクが別のバイクに衝突する事故があった。

事故があったのは、茅ヶ崎市の国道の交差点で、直進してきた2人乗りのバイクが右折しようとしたバイクと衝突した。

この事故で、2人乗りのバイクに一緒に乗っていた女子大学生（18）が全身を強く打つなどしその場で死亡が確認された。

福岡空港で、女性利用客が「保安検査を受けずに待合室に入った」と申し出たため、再検査などで約2千人に影響が出た。

しかし、約100人が再検査を終えたところで、この女性が実際には保安検査場を通過していたことが判明した。

女性は出発前の名古屋行きジェットスター便に搭乗予定で、当時は酒に酔っていたという。

野沢温泉村で昨夜民家の外壁などが焼け、警察は、放火の可能性もあるとみて調べている。

野沢温泉村豊郷の住宅から火が出て1階の外壁と鳥小屋が焼けた。同じ地区では、4日前にも現場から300メートルほど

離れた民宿で客室が焼けたほか、去年も原因が分からない火事が続き、さらに4件の不審火があった。

高知・土佐市の海で、行方不明になっていた小学6年の男の子とみられる遺体が、15日夕方に見つかった。

午後5時半ごろ通報があり警察が駆けつけたところ、身長130cmほどの男の子とみられる遺体が見つかった。

先月8日から行方不明になっている小学6年の少年(11)と服装などが似ており、警察が身元の確認を急いでいる。

久留米市で、５歳の男の子がクリークに浮いているのが見つかり、その後死亡した。

久留米市三潴町で、「父親が買い物に行った隙に息子が家を出て行った」と男の子の母親から１１０番通報があった。

通報から約４０分後、男の子（５）は自宅近くのクリークで浮いている状態で発見され、搬送先の病院で死亡が確認された。

山形発 名古屋空港行きの　　　　便は、離陸に向けて走行中に滑走路から逸脱して芝生の上で停止した。

乗客・乗員合わせて64人にケガはなかったが、国は事故につながりかねない重大インシデントと認めた。

運輸安全委員会の航空事故調査官による現地調査を始めました。

容疑者(55)は、同居する元夫とみられる男性を鉄製のハンマーで殴り、熱湯をかけて殺害した疑いが持たれている。

容疑者は、「元夫がひっくり返って動けない」と、自ら119番通報していたことがわかった。

男性が床で寝ていたため、ベッドに行くよう言ったものの、動かなかったことに腹を立てて犯行に及んだという。

合成麻薬「MDMA」を学生に作らせたり、別の麻薬を所持したりしたとして、松山大薬学部教授(61)が書類送検された。

医療薬学科の教授で、危険ドラッグなどの薬物を研究していたという。

麻薬研究者の免許がないにも関わらず、学生に作らせた理由について「勉強のためにやらせていた」と話している。

岡山県美作市で女性(96)の遺体が自宅付近で見つかった事件で、近くに住む男(84)を殺人容疑で再逮捕した。

容疑者の自宅北側の舗装されていない路上で、被害者の頭部などを鈍器のようなもので複数回殴り、殺害した疑い。

容疑者はすでに死体遺棄容疑で逮捕されていたが、「殺してはいません」と殺人容疑を否認しているという。

宮崎市の住宅街で25日午前、、縦横高さ約30 X 40 X 20センチの不審な段ボール箱が発見された。

箱はカタカナでサリンと書かれた紙がで貼られており、警察消防などが駆け付け現場は一時騒然となった。

警察官は周辺住民に対して窓を閉め、屋外に出ないように呼び掛けた。

佐賀市の結婚式場に侵入し窓ガラスや椅子などを壊したとして、24歳の男子大学生が建造物損壊の疑いで逮捕された。

男は犯行の数日前に結婚式最中の同式場に怒鳴り込んでいた。

男は容疑を認めており、式場には鉄製のパイプが落ちていたという。

福岡県警博多署は20日、タクシー運転手の男(64)を道交法違反容疑で逮捕した。

缶ビールを手に車から降りてきた男を見た男性が110番し、飲酒検知で基準値の倍近いアルコール分が検出された。

男は、酒は昼頃に飲んだからアルコール分が出るはずがないと否認しているという。

県教委は29日、奄美市内で酒気帯び運転をしたとして、県立学校に勤務する教諭(45)を停職6か月の懲戒処分とした。

教諭は自宅で焼酎の水割りを飲んだ後に乗用車を運転。シートベルトを着用していなかったことから職務質問を受け、

酒気帯び運転が発覚した。教諭は「取り返しのつかないことをしてしまい、反省している」と話しているという。

吉井川に設置されている北浜水門が何者かによって開けられ、海水が水田に流れ込んでいるのを住民が見つけた。

海水が流れ込んだのは岡山市東区正儀の水田で、浸水した農地は約16ヘクタールに上る。

現場では、急ピッチで水田から塩分を取り除く作業が進められている。

2. NEWS STIMULI (NEUTRAL)

平成の終わりとともに再び注目を集めている岐阜県関市の平成地区で新たな「平成」グッズが発売さた。

22日に発売されたのは、その名も「平成の空気」。平成時代の空気を残そうと企画され、

関市平成地区にある「元号橋」の上で採取された。缶詰の中には、平成に製造された5円玉も入ってる。

NHKの[redacted]で衣装などの監修を務めた伊那市出身の[redacted]夫さんの作品展が開かれている。

この作品展は、ドラマや映画の登場人物の衣装やヘアメーク、

それに小道具などの監修を務める[redacted]さんの世界を知ってもらおうと、伊那市の信州高遠美術館が企画した。

ドライフラワーと固まる樹脂でつくるオブジェの体験教室が、弘前市で開かれた。

体験教室は、地元でオブジェ作りを続けている女性が開き、会場には、４人が集まった。

はじめに参加した人たちはおよそ２０種類のドライフラワーから、気に入ったものを選んでいった。

釜石市では、ご当地グルメの「釜石ラーメン」を提供する30店を日本語と英語で紹介する地図が作成され、

国際発信の取り組みの１つとして注目されている。マップには、定番ラーメンの写真と店の情報が詳しく紹介されている。

このマップは、7,000部発行され、釜石駅近くの観光総合案内所や市内のホテル、それにJR新花巻駅でも配布される。

村山市の最上川美術館でさまざまな布を縫い合わせて作ったキルト作品の展示会が開かれている。

村山市内のキルト教室で学ぶ愛好家たちで作るグループが開き、会場には、壁掛けやバッグなどの作品が並べられている。

縦横ともおよそ2mある大きな壁掛けは、絹の小さな布を１針ずつ手作業で縫い合わせ、完成まで3年かかった。

[redacted]株式会社は、15.6型ノート[redacted]製品を発売した。

[redacted]の後継機となる15.6型クラムシェルノートPC。

第8世代[redacted]プロセッサを搭載し、M.2 SSDを追加。M.2 SSD選択時は、最大2基のストレージが搭載可能となっている。

[redacted]の年次イベント[redacted]で人気ビデオゲーム手がける[redacted]での[redacted]活用が披露された。

[redacted]は[redacted]などの人気選手が登場する、ゲーム市場もっとも売れたビデオゲーム[redacted]を開発している。

その開発からローンチ、運用は、600人が関わる一大プロジェクトであり、コラボレーションツールの[redacted]が大活躍したという。

[redacted]会社は、[redacted]サービスの元社長[redacted]新社長に迎える方針を固めたことがわかった。

海外経験が豊富なことから国際線など誘致に手腕を期待されての就任とみられ、6月の株主総会を経て正式決定する。

[redacted]会社のトップ交代は4年ぶりで、同社社長にトヨタ出身者が就くのは会社の設立から5代連続。

[redacted]は25日、10月1日付で持ち株会社体制に移行し、社名を[redacted]に変更すると発表した。

ネット関連事業を手掛ける[redacted]と、金融事業を統括する中間持ち株会社をそれぞれ設立し、持ち株会社の傘下に収める。

[redacted]と共同出資する[redacted]を中心とする金融事業の強化と、より迅速な事業運営体制の構築が狙いだ。

[redacted]は10日、2020年1月1日をめどに純粋持ち株会社に移行する方針を決めたと発表した。

100%出資の子会社を設立して国内事業を移管し、海外事業の統括会社と並んで持ち株会社の傘下に置く計画だ。

国内と海外の両方に目配りしやすい体制を構築するほか、国内事業の意思決定を迅速化する狙いがある。

[redacted]は、家族連れの訪日外国人をターゲットにしたホテルの宿泊事業に参入すると発表した。

1泊、1人あたり4,000円から宿泊が可能で、1LDKの客室にすることで「都市型のコンドミニアム」を提案するとしている。

ホテル内には、[redacted]製の美容家電が並ぶ専用スペースを設置し、製品の販売拡大にもつなげたい考え。

[redacted]自動車は、研究開発のスピードを加速させるため、約3000億円を投資してテストコースを新たに整備している。

豊田市の本社の近く、豊田市と岡崎市にまたがる山間部の一帯、650ヘクタールの敷地に研究開発施設の建設を進める。

25日は、その一部にあたる全長約5.3キロのテストコースが完成し、関係者およそ70人が出席して竣工式が行われた。

[redacted]は、海外事業を統括する子会社をシンガポールに設立すると発表した。

中間持ち株会社としてシンガポール、ハワイなど海外の事業会社の管理を担う。

社長には[redacted]創業者の[redacted]氏が就任する。

自動車企業の米[redacted]は22日、完全自動運転車技術を紹介するプライベートイベント[redacted]を開催し、

同社が開発したコンピュータおよびプロセッサを公開した。車載コンピュータの形状は従来とほぼ同等だが、

プロセッサは[redacted]から[redacted]自社開発のものへと置き換わった。

[redacted]の資産運用部門[redacted]グループは、事業規模の拡大を模索している。予備的段階ではあるが、

スイスの[redacted]グループや仏[redacted]、仏[redacted]といった企業と提携を協議している。

合弁事業などを含むあらゆる選択肢を検討しているが、[redacted]銀行が経営権を手放すシナリオは想定してないという。

[redacted]は4月25日、[redacted]とマネジメント契約を締結したと発表した。

[redacted]のゲーミングチームとして結成されたeスポーツチーム。

同社はIT関連事業のノウハウを生かし、[redacted]のマネジメント事業を通じて、eスポーツ市場の拡大に努めるとしている。

富山駅北と岩瀬浜を結ぶ富山[redacted]が、富山[redacted]に吸収合併されることが発表された。

富山駅では、来年3月に路面電車が高架下でつながるを控えていて、合併は、その前の来年2月を予定している。

富山[redacted]によると南北接続後に、2つの会社で同じ運賃での運行が難しくなることなどから、合併することとなった。

アメリカ発の小型スマホが、日本に上陸した。小ささすぎるスマホ[redacted]。縦9.6cm、横5cm、重さは、なんと62.5グラム。

ネット、通話、写真などの機能は、一般的なスマートフォンとほぼ変わらない。

大型化が進むスマホ市場に一石を投じるこのスマホは、全国の家電量販店で24日から販売開始。。

[redacted]は、絶大な人気を誇る[redacted]のアップデート版をリリースした。

この[redacted]拡張機能は、[redacted]の[redacted]のマーケットプレースからダウンロード可能だ。

[redacted]拡張機能は現在、オンラインストアから800万回以上インストールされている。

水戸市中心部のアリーナの名称が、市内のソフトウェア開発会社の名前をとって[redacted]に決まった。

同アリーナはバスケットボール[redacted]のイベントなどに使われている。

[redacted]の[redacted]社長と、[redacted]の[redacted]社長などが記者会見して明らかにした。

[redacted]が飛行機に携帯電話の基地局を載せて成層圏を飛ばす事業に乗り出すことが24日、分かった。

本格導入を控える第5世代（5G）移動通信システムでの活用を見込む。

投資額を抑えながら広大な通信エリアを実現する狙い。

新型[redacted]のメインカメラの画素数が1200万画素にアップグレードされると著名アナリストが予測。

[redacted]が今年発売する新型[redacted]は、前面カメラが現行の700万画素から1200万画素にアップグレードする。

著名アナリストの[redacted]氏による予測を、米メディア[redacted]が4月18日に伝えた。

[redacted]の3社が去年５月から開始したメッセージアプリ「＋メッセージ」の機能強化を発表した。

このメッセージアプリのサービスは[redacted]に対抗するため、携帯電話大手3社が去年から始めたもの。

新機能では企業の公式アカウントとのやり取りが可能となった。

[redacted]は、「2019年 国内IoTエッジインフラストラクチャの選定基準調査」の結果を発表した。

IoTエッジインフラベンダー選定では、「IoTの技術力」と「ユーザー業務の理解度」が優先されることが分かった。

この調査は、2018年12月にアンケート形式で実施され、509の組織から得られた回答をもとに分析している。

埼玉県小川町は、町商工会や町内の酒蔵と連携して「小川町3蔵酒蔵めぐり」を開催した。

酒蔵では地酒や甘酒がふるまわれた他、周辺では地元飲食店による出店も開かれ、町内外から来た人でにぎわった。

町には三つの酒蔵があり、それぞれこだわりの日本酒を醸造、販売している。

6月にG20が開催される大阪で会場となる大阪市の[redacted]を警察庁長官（60）が会場を視察した。

6月28日と29日に開かれる「G20大阪サミット」に伴い、大阪では阪神高速が封鎖されるなど大規模な警備態勢が敷かれる。

[redacted]長官は「全国の警察で大阪府警を支援し、警備を全うしたい」と話した。

3. NEWS STIMULI (POSITIVE)

長崎市の商店街を抜けた先で、川の護岸に埋め込まれたはしごの上で震えている猫が見つかった。

気づいた買い物客らが集まるなか、傘や買い物かごを持った男性が登場した。

猫を買い物かごに入れようとしたが入れられず、猫を右腕で抱え、左手だけではしごをのぼり、無事助け出した。

佐賀市消防局は、嘉瀬川で溺れていた母子３人を救助した男性会社員と病院事務の男性２人に感謝状を贈った。

男性会社員は大相撲の元力士で、急流の川で３人を抱えて踏ん張った。

病院事務の男性は電線の延長コードを川に投げ、周囲の協力も得て全員を引き上げた。

清水海上保安部は、海に転落した釣り人を救助した富士市の遊漁船長の男性（39）に感謝状を贈った。

男性は、富士市の浦港を船で出港する時に防波堤近くで釣りをしていた70代男性が海に転落するのを目撃した。

その直後に自身の操船で現場へ急行し、救命浮輪を使って男性を無事救助した。

沼津署は、事故現場などで保護活動に協力した高校生たちに署長感謝状を贈った。

表彰されたのは桐陽高校の男子高校生４名だった。

彼らは橋から狩野川に転落しそうになっていた女性を発見し、警察に通報し、女性の手を支えるなど救助に貢献した。

沖縄県南城市にあるゴルフ倶楽部で79歳の男性がプレー中に倒れ心肺停止状態になった。

ゴルフ倶楽部のキャディーの女性が、男性にAEDや心臓マッサージを施して一命を救った。

女性は「退院したと聞き安心した。多くの人にAEDの使い方を覚えてもらい、人命救助につながれば」と話した。

静岡市消防局は静岡マラソンで人命救助を行った国士舘大モバイルAED隊と、男性医師（38）に感謝状を贈呈した。

モバイルAED隊がゴール直後に倒れた男性と37.2キロ地点で倒れた男性を発見し、心肺蘇生処置を施した。

37・2キロ地点で倒れた男性の救命には近くを走っていた男性医師も協力した。

人命救助に貢献したとして深谷署は、本庄市の大学准教授の男性（42）に感謝状を贈呈した。

同署によると、男性が市道を乗用車で走行中、行方不明者に似ている高齢男性（82）を見つけて警察に連絡、

車中で安全確保に務め、人命救助に貢献した。男性は「今後も協力をしていきたい」と話していた。

東京消防庁はとっさの応急手当で70代の男性客の命を救ったとして、男性店長に感謝状を贈った。

家族4人で店を訪れた70代の男性が突如苦しみ始めた。

店長はステーキを喉に詰まらせて苦しむ男性の元に駆け寄り、冷静な対応で男性の命を取り留めた。

人命救助に貢献したとして、鉾田市新鉾田の自営業の夫妻に、鉾田消防署から感謝状が贈られた。

夫妻は自宅近くの川沿いを飼い犬の「モモ」(雌6歳)を連れて散歩中、倒れていた高齢女性を助けたもの。

夫妻は発見のきっかけがモモだったことから、「モモのお手柄。ご褒美をやりたい」と笑顔を見せた。

兵庫県警垂水署はこのほど、同区に住む少年3人に同県の善行賞「のじぎく賞」を贈った。

道端で動けなくなっていた女性(86)を助け、自宅まで送り届けた。

3人は小学校から一緒にサッカーをする大の仲良しで、連携した好プレーが人命救助につながった。

朝霞署は、和光市の「朝日新聞販売所和光市店」社員(44)に独居の高齢男性の命を救ったとして感謝状を贈呈した。

同社員は郵便受けにたまった新聞を見て異常を察知し、自らの休日に訪ねるなどして通報警察に通報した。

署長は「休みの日に再訪することは、なかなかできることではない」と述べた。

川口市江戸袋2丁目の運送会社事務室で発生した火災で、南消防署は隣にある鋳物工場の従業員3人に感謝状を贈った。

3人が初期消火活動を行い、3階建て建物の1階部分の被害に食い止めることができた。

「3人の迅速かつ適切な初期消火活動がなかったら大火災になった」と消防署長はコメントしている。

浦和西署は、強制わいせつ事件で容疑者の逮捕に貢献したとして、さいたま市の専門学校生(18)に感謝状を贈呈した。

被害者の女性が助けを求める声を聞き、即座に容疑者を追い掛けて確保した。

「何が何だか分からなくて、取りあえず追い掛けようと体が勝手に動いた。人の役に立ててよかった」と控えめに喜んだ。

岡山県は西日本豪雨のときに浸水した倉敷市真備町で、家に取り残された住民らを助けた市民に感謝状を贈った。

同町・市場の男性は、有井の男性からの救助要請に自分の釣り船で駆けつけ、男性の家族を助けた。

その後に助けた男性の長男らと浸水した家に残された住民らを助けた。救助活動は約8時間に50人ほどを避難させた。

町田市在住の男性(72)が、大石山のロッジで、滑落して行方不明になっていた小田原市の男性(37)を発見し、救助した。

遭難した男性は滑落後、約8日間沢の水で生き延び、発見当時は足の骨を折る大ケガをしていたが一命をとりとめている。

男性は遭難者を発見後、服と食料を与えたうえで携帯電話がつながる場所まで一人で山を越えて行き救助を求めた。

会津若松市の小学生の兄弟が排水路に倒れていた男性を救助し、会津若松署から感謝状が贈られた。

2人は下校中に男性を発見して自宅にいた母親(46)に報告した。現場に赴いた母親は近隣住民に119番通報を頼みんだ。

兄弟は両親らと男性を救助した。男性は同市の病院に搬送され、命に別条はなかった。

約30年にわたり献血を続けてきた福井県福井市の男性(69)が長年の功績をたたえられ感謝状を贈られた。

県赤十字血液センターによると、30年に219回は県内最多ではないが、ほとんどいない、とのこと。

男性は30代のころ献血を始め、年10回ペースで血液を提供してきた。「続けることができて良かった」と笑顔を見せた。

24日午前7時半過ぎ、神奈川県小田原市の畑で、近くに住む92歳の男性がオスのイノシシに襲われた。

男性はイノシシに両手両足をかまれたが、つえやクワで撃退したという。

その後、病院に搬送されたが命に別条はないという。

米 は4月13日、世界最大の飛行機 の初飛行に成功した。

同機は、空中からのロケットの発射を目的として開発されている飛行機で翼長は117m、6基のエンジンを搭載している。

公式発表によれば、カリフォルニア州のモハベ空港にて実施した2.5時間の飛行では高度5180m、時速304kmに到達した。

アメリカの宇宙開発企業 の超大型ロケット「ファルコンヘビー」が、初の商業打ち上げを成功させた。

同機はケネディ宇宙センターから打ち上げられ、搭載していた通信衛星を切り離し、初の商業打ち上げを成功させた。

ファルコンヘビーは、ブースターが再利用でき、今回、3つのブースターすべてが予定された場所に帰還した。

「世界一髪の長い10代」としてギネス世界記録に認定された鹿児島県出水市の女性(18)が黒髪に初めてはさみを入れた。

切り取った1メートル超の髪の毛は、医療用かつらを扱うNPO法人に寄付する。

「私の髪で少しでも明るい気持ちになってもらえれば」と、半分以上をカットした。

伊万里・有田消防本部は共同住宅で発生した火災で、住民を救助した東山代町在住の男性(36)に感謝状を贈った。

男性が叫び声に気付いてベランダに出ると、隣の住居から煙が出ており、足の不自由な一人の女性が助けを求めていた。

男性はベランダの隔て板を蹴破って女性を救出し、二つ隣の住居の高齢者も助け出した。

イスラエル・テルアビブ大学の研究チームが、人間の細胞や生体物質を使った人工心臓を3Dプリンターで試作したと発表した。

同大によれば3Dプリンターで血管や心房、心室などが備わった人工心臓が作られたのは「世界で初めて」という。

実用化に向けてなお開発が必要だが、将来の心疾患治療への応用が期待される。

山梨大学の研究グループがドローンと先端技術を活用した野性のニホンジカの捕獲に国内で初めて成功した。

野性のニホンジカは県内で約5万頭余り。木を食い荒らすなど生態系を保護する上で課題となっている。

適正な生息数のおよそ11倍で増え過ぎたシカの効率的な捕獲につながると期待されている。

アラブ首長国連邦の女性が、27年ぶりに昏睡（こんすい）状態から目覚め、回復しつつある。

女性は車で学校に息子を迎えに行った際、バスと衝突し、脳に深刻な損傷を負った。

昏睡状態はその後27年間続いたが、ドイツの病院で昨年、意識を回復した。

千グラム未満の超低出生体重児として生まれた那覇市の男子双子（7）がこの春、1年遅れで小学校に入学する。

新生児集中治療室（NICU）で長期間を過ごし、大手術も経験した兄弟は家族で琉球大学付属病院を訪れ、

出生時から2人を見守ってきたスタッフと再会し、元気いっぱいに入学を報告した。

札幌市豊平区の豊平公園にある野草園で、「春の妖精」と呼ばれるカタクリの花が満開になった。

園内には約80平方メートルのカタクリの群落があり、紫がかったピンク色の花びらが、うつむきかげんに咲く。

神奈川県から友人3人と訪れた男性（70）は「街の真ん中で見られるとは、さすがは北海道」と感心していた。

伊東市の伊豆シャボテン動物公園で、動物たちが出産期を迎え、次々にかわいらしい赤ちゃんが誕生している。

アフリカ南部に生息する、二本足で立つ姿が人気のミーアキャット。4匹の赤ちゃんが誕生し、順調に育っている。

誕生からひと月余りで体長は約15センチに。巣箱から出て地面を掘ったり、兄弟とじゃれあったりしている。

岐阜県郡上市の「國田家の芝桜」が見頃を迎えた。シバザクラが絨毯のように斜面に広がり、奥美濃の春を彩っている。

シバザクラは、故國田さんが自宅裏で株分けし増やしてきた。彼女が亡くなった後も家族と住民有志が手入れしている。

今年は4月に入って積雪を記録するなど肌寒い日が続き、昨年より10日ほど遅く開花。

[redacted]電鉄伊太祈曽駅で、特産の南高梅をイメージした電車の車両に「令和」を記念したヘッドマークが取り付けられた。

記念行事には駅長を務める三毛猫「よんたま」も出席し、和歌山、梅にも縁深い新元号の決定を祝った。

令和が万葉集の「梅花の歌」から選定されたことにちなみ、電車と「祝・令和」の文字とともに、梅の花やうぐいすが描かれた。

The names of commercial companies and actual persons appeared in the news have been blacked out.

REFERENCES

Albraikan, A., Hafidh, B., & Saddik, A. E. (2018). iAware: A real-time emotional biofeedback system based on physiological signals. *IEEE Access*, *6*, 78780-78789.

Amano, S., & Hirata, Y. (2010). Perception and production boundaries between single and geminate stops in Japanese. *Journal of the Acoustical Society of America*, *128*(4), 2049-2058.

Anderson, J. E., & Holcomb, P. J. (1995). Auditory and visual semantic priming using different stimulus onset asynchronies: An event-related brain potential study. *Psychophysiology*, *32*(2), 177-190.

Antelmi, I., de Paula, R. S., Shinzato, A. R., Peres, C. A., Mansur, A. J., & Grupi, C. J. (2004). Influence of age, gender, body mass index, and functional capacity on heart rate variability in a cohort of subjects without heart disease. *The American Journal of Cardiology*, *93*(3), 381-385.

Ardila, A., Bernal, B., & Rosselli, M. (2014). Participation of the insula in language revisited: a meta-analytic connectivity study. *Journal of Neurolinguistics*, *29*, 31-41.

Azevedol, R.T., Bennett, N., Bilicki, A., Hooper, J., Markopoulou, F., & Tsakiris, M. (2017). The calming effect of a new wearable device during the anticipation of public speech. *Scientific Reports*, *7*(2285).

Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences*, *22*(4), 577-609.

Barsalou, L. W. (2008). Grounded cognition. *Annual Review of Psychology*, *59*, 617-645.

Barsalou, L. W. (2010). Grounded cognition: past, present, and future. *Topics in Cognitive Science*, *2*(4), 716-724.

Barsalou, L. W., Santos, A., Simmons, W. K., & Wilson, C. D. (2008). Language and simulation in conceptual processing. In M. de Vega, A. M. Glenberg & A. C. Graesser (Eds.), *Symbols, embodiment, and meaning: debates on meaning and cognition* (pp. 245-283). Oxford University Press.

Bennett, M. R., & Hacker P. M. (2005). Emotion and cortical-sub cortical function: conceptual developments. *Progress in Neurobiology*, *75*(1), 29-52.

Bergen, B. K. (2019). Embodiment. In Dąbrowska, E., & Divjak, D. (Eds.), *Cognitive Linguistics - Foundations of Language* (pp. 11-35). De Gruyter Mouton.

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

Bergen, B. K., Lindsay, S., Matlock, T., & Narayanan, S. (2007). Spatial and linguistic aspects of visual imagery in sentence comprehension. *Cognitive Science*, *31*(5), 733-764.

Bergman, P., Ho, H.-N., Koizumi, A., Tajadura-Jiménez, A., & Kitagawa, N. (2015). The pleasant heat? Evidence for thermal-emotional implicit associations occurring with semantic and physical thermal stimulation. *Cognitive Neuroscience*, *6*(1), 24-30.

Borhani, K., Làdavas, E., Fotopoulou, A., & Haggard, P. (2017). "Lacking warmth": Alexithymia trait is related to warm specific thermal somatosensory processing. *Biological Psychology*, *128*, 132-140.

Bower, G. H., & Morrow, D. G. (1990). Mental Models in Narrative Comprehension. *Science*, *247*(4938), 44-48.

Bradley, M. M., & Lang, P. J. (1994). Measuring emotion: The Self-Assessment Manikin and the semantic differential. Journal of Behavior *Therapy and Experimental Psychiatry*, *25*(1), 49-59.

Brancazio, L., Miller, J.L., Paré, M.A. (2003). Visual influences on the internal structure of phonetic categories. *Perception & Psychophysics*, *65*, 591-601.

Bransford, J. D., Barclay, J. R., & Franks, J. J. (1972). Sentence memory: A constructive versus interpretive approach. *Cognitive Psychology*, *3*(2), 193-209.

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.

Buccino, G., Riggio, L., Melli, G., Binkofski, F., Gallese, V., & Rizzolatti, G. (2005). Listening to action related sentences modulates the activity of the motor system: A combined TMS and behavioral study. *Cognitive Brain Research*, *24*(3), 355-363.

Buchwald, A. B., Winters, S.J., Pisoni, D.B. (2009). Visual speech primes open-set recognition of spoken words, *Language and Cognitive Processes*, *24*(4), 580-610,

Burgess, C., & Lund, K. (1997). Modelling parsing constraints with high-dimensional context space. *Language and Cognitive Processes*, *12*(2-3), 177-210.

Chabris, C. F., Heck, P. R., Mandart, J., Benjamin, D. J., & Simons, D. J. (2019). No evidence that experiencing physical warmth promotes interpersonal warmth. *Social Psychology*, *50*(2), 127-132.

Chapman, S. B., Aslan, S., Spence, J. S., DeFina, L. F., Keebler, M. W., Didehbani, N., & Lu, H. (2013). Shorter term aerobic exercise improves brain, cognition, and cardiovascular fitness in aging. *Frontiers in Aging Neuroscience*, *5,* Article 75.

Chen, Y.-C., & Spence, C. (2011). Crossmodal semantic priming by naturalistic sounds and spoken words enhances visual sensitivity. *Journal of Experimental Psychology: Human Perception and Performance*, *37*(5), 1554-1568.

Chen, Z., Poon, K.-T., & DeWall, N. (2014). Cold thermal temperature threatens belonging: the moderating role of perceived social support. *Social Psychological and Personality Science*, *6*(4), 439-446.

Chow, H.M., Mar, R.A., Xu,Y. Liu,S., Wagage, & Braun, A.R. (2015). Personal experience with narrated events modulates functional connectivity within visual and motor systems during story comprehension. *Human Brain Mapping*, *36*(4), 1494-1505.

Citron, F. M. M., & Goldberg, A. E. (2014). Social context modulates the effect of physical warmth on perceived interpersonal kindness: a study of embodied metaphors. *Language and Cognition*, *6*(1), 1-11.

Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review*, *82*(6), 407-428.

Connell, L., & Lynott, D. (2009). Is a bear white in the woods? Parallel representation of implied object color during language comprehension. *Psychonomic Bulletin & Review*, *16*, 573-577.

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.

Costa, J., Adams, A.T., Jung, M. F., Guimbeti, F., & Choudhury, T. (2016). Emotion Check: Leveraging bodily signals and false feedback to regulate our emotions. *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, 758-769.

Craig, A. (2002). "How do you feel? Interoception": the sense of the physiological condition of the body. *Nature Reviews Neuroscience*, *3*(8), 655-666.

Craig, A. D. (2004). Human feelings: why are some more aware than others? *Trends in Cognitive Science*, *8*(6), 239-41.

Critchley, H. D., & Garfinkel, S. N. (2018). The influence of physiological signals on cognition. *Current Opinion in Behavioral Sciences*, *19*, 13-18.

Critchley, H. D., Corfield, D. R., Chandler, M. P., Mathias, C. J., & Dolan, R. J. (2000). Cerebral correlates of autonomic cardiovascular arousal: a functional neuroimaging investigation in humans. *The Journal of physiology*, *523*(1), 259-270.

Critchley, H. D., Eccles, J., & Garfinkel, S. N. (2013). Interaction between cognition, emotion, and the autonomic nervous system. In R. M. Buijs & D. F. Swaab (Eds.), *Handbook of Clinical Neurology* (pp. 59-77). Elsevier.

Critchley, H. D., Wiens, S., Rotshtein, P., Ohman, A., Dolan, R.J. (2004). Neural systems supporting interoceptive awareness. *Nature Neuroscience*, *7*(2),189-195.

Damasio, A. R. (1994). *Descartes' error: emotion, reason, and the human brain*. New York: Grosset/Putnam.

Damasio, A. R. (1998). Emotion in the perspective of an integrated nervous system. *Brain Research Reviews*, *26*(2-3), 83-86.

Damasio, A. R., & Tranel, D. (1993). Nouns and verbs are retrieved with differently distributed neural systems. *Proceedings of the National Academy of Sciences of the United States of America*, *90*(11), 4957-4960.

Damasio, A., Grabowski, T., Bechara, A., Damasio, H., Ponto, L. L. B, Parvizi, J, & Hichwa, R. D. (2000). Subcortical and cortical brain activity during the feeling of self-generated emotions. *Nature Neuroscience*, *3*, 1049-1056.

D'Ausilio, A., Pulvermüller, F., Salmas, P., Bufalari, I., Begliomini, C., & Fadiga, L. (2009). The Motor Somatotopy of Speech Perception. *Current Biology*, *19*(5), 381-385.

Dijkstra, T., & van Heuven, W. J. B. (2002). The architecture of the bilingual word recognition system: from identification to decision. Bilingualism: *Language and Cognition*, *5(3)*, 175-197.

de Koning, B.B., Stephanie I. Wassenburg, Lisanne T. Bos, & Menno Van der Schoot (2017). Size Does Matter: Implied Object Size is Mentally Simulated During Language Comprehension. *Discourse Processes*, *54*(7), 493-503.

De Saussure, F. (1959). *Course in general linguistics*. (W. Baskin, Trans.) New York: McGraw-Hill.

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.

de Vignemont, F., & Massin, O. (2015). Touch. In M. Matthen (Ed.), *The Oxford Handbook of Philosophy of Perception* (pp. 294-313). Oxford University Press.

Erber, N. P. (1969). Interaction of audition and vision in the recognition of oral speech stimuli. *Journal of Speech, Language, and Hearing Research*, *12*(2), 423-425.

Ernst, G. (2017). Heart-Rate Variability—More than Heart Beats? *Frontiers in Public Health*, *5,* Article 240.

Estes, Z., Verges, M., & Barsalou, L. W. (2008). “Head Up, Foot Down”: Object Words Orient Attention to the Objects’ Typical Location. *Psychological Science*, *19*(2), 93-97.

Fadiga, L., Craighero, L., Buccino, G., & Rizzolatti, G. (2002). Speech listening specifically modulates the excitability of tongue muscles: A TMS study. *European Journal of Neuroscience*, *15*(2), 399-402.

Fetterman, A. K., Wilkowski, B. M., & Robinson, M. D. (2018). On feeling warm and being warm: Daily perceptions of physical warmth fluctuate with interpersonal warmth. *Social Psychological and Personality Science*, *9*(5), 560–567.

Fischer, M. H., & Zwaan, R. A. (2008). Embodied language: A review of the role of the motor system in language comprehension. *The Quarterly Journal of Experimental Psychology*, *61*(6), 825-850.

Fodor, J. A. (1998). *Concepts: Where Cognitive Science Went Wrong*. Oxford: Oxford University Press.

Foltz, P. W., Kintsch, W., & Landauer, T. K. (1998). The measurement of textual coherence with latent semantic analysis. *Discourse Processes*, *25*(2-3), 285-307.

Galantucci, B., Fowler, C., & Turvey, M. T. (2006). The motor theory of speech perception reviewed. *Pyschonomic Bulletin & Review*, *13*, 361-377.

Gallace, A., & Spence, C. (2014). *Touch with the Future: The sense of touch from cognitive neuroscience to virtual reality*. Oxford: Oxford University Press.

Gallese, V., Fadiga, L. Fogassi, L., & Rizzolatti, G. (1996). Action recognition in the premotor cortex. *Brain*, *119*(2), 593-609.

Greenspan, J. D., Roy, E. A., Caldwell, P. A., and Farooq, N. S. (2003). Thermosensory intensity and affect throughout the perceptible range. *Somatosensory & Motor Research*, *20*(1), 19-26.

Gentilucci, M., Benuzzi, F., Bertolani, L., Daprati, E., & Gangitano, M. (2000). Language and motor control. *Experimental Brain Research*, *133*(4), 468-490.

Gick, B., Derrick, D. (2009). Aero-tactile integration in speech perception. *Nature*, *462*(7272), 502-504.

Glenberg, A. M., & Kaschak, M. P. (2003). The body's contribution to language. In Ross, B. H. (Ed), *The Psychology of Learning and Motivation.* (pp. 93-126). Academic Press.

Glenberg, A. M., & Robertson, D. A. (1999). Indexical understanding of instructions. *Discourse Processes*, *28*(1), 1-26.

Glenberg, A. M., & Robertson, D. A. (2000) Symbol grounding and meaning: A comparison of high-dimensional and embodied theories of meaning. *Journal of Memory and Language*, *43*(3), 379-401.

Glenberg, A. M., Meyer, M., & Lindem, K. (1987). Mental models contribute to foregrounding during text comprehension. *Journal of Memory and Language*, *26*(1), 69-83.

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.

Glenberg, A. M., Webster, B. J., Mouilso, E., Havas, D., & Lindeman, L. M. (2009). Gender, emotion, and the embodiment of language comprehension. *Emotion Review*, *1*(2), 151-161.

Glenberg, A. M., Witt, J. K., Metcalfe, J. (2013). From the revolution to embodiment: 25 years of cognitive psychology. *Perspectives on Psychological Science*, *8*(5), 573-585.

Goldinger, S. D., Luce, P. A., Pisoni, D. B., Marcario, J. K. (1992). Form-based priming in spoken word recognition: The roles of competition and bias. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*(6), 1211-1238.

Grabe, M. E., Lang, A., & Zhao, X. (2003). News content and form: Implications for memory. *Communication Research*, *30*(4), 387-413.

Grainger J., & Ferrand, L. (1994). Phonology and orthography in visual word recognition: Effects of masked homophone primes. *Journal of Memory and Language, 33*(2), 218-233.

Grainger, J., Diependaele, K., Spinelli, E., Ferrand, L., & Farioli, F. (2003). Masked repetition and phonological priming within and across modalities. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 29(6), 1256-1269.

Habermas, T., & V. Diehl. (2013). The episodicity of verbal reports of personally significant autobiographical memories: Vividness correlates with narrative text quality more than with detailedness or memory specificity. *Frontiers in Behavioural Neuroscience, 7,* Article 110.

Halali, E., Meiran,N., & Shalev, I. (2017). Keep it cool: temperature priming effect on cognitive control. *Psychological Research*, *81*(2), 343–354.

Hallman, D. M., Holtermann, A., Søgaard, K., Krustrup, P., Kristiansen, J., & Korshøj, M. (2017). Effect of an aerobic exercise intervention on cardiac autonomic regulation: A worksite RCT among cleaners. *Physiology & Behavior*, *16*9, 90-97.

Han, M. (1994). Acoustic manifestations of mora timing in Japanese. *Journal of the Acoustical Society of America*, *96*(1), 73-82.

Hauk, O., Shtyrov, Y., & Pulvermüller, F. (2008). The time course of action and action-word comprehension in the human brain as revealed by neurophysiology. *Journal of physiology-Paris, 102*(1-3), 50-58.

Havas, D. A., Glenberg, A. M., & Rinck, M. (2007). Emotion simulation during language comprehension. *Psychonomic Bulletin & Review*, *14*(3), 436-441.

Hecht, D., & Reiner, M. (2010). Stroop interference and facilitation effects in kinesthetic and haptic tasks. *Advances in Human-Computer Interaction, 2010,* 852420.

Hirata, Y., & Whiton, J. (2005) Effects of speech rate on the singleton/geminate distinction in Japanese. *Journal of the Acoustical Society of America*, *118*(3), 1647-1660.

Ho, H. N., Van Doorn, G. H., Kawabe, T., Watanabe, J., & Spence, C. (2014). Colour-temperature correspondences: when reactions to thermal stimuli are influenced by colour. *PLoS ONE*. *9(3),* e91854.

Ho, H.-N., & Jones, L.A. (2004). Material identification using real and simulated thermal cues*. Proceedings of 26th Annual International Conference of the IEEE Engineering in Medicine and Biology Society,* 2462-2465.

Ho, H.-N., & Jones, L.A. (2006). Thermal model for hand-object interactions. *IEEE Proceedings of the 14th Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems*, 461-467.

Ho, H.-N., & Jones, L.A. (2007). Development and evaluation of a thermal display for material Identification. *ACM Transactions on Applied Perception*, *4*(2), 1-24.

Ho, H.-N., & Jones, L.A. (2008). Modeling the thermal responses of the skin surface during hand-object interactions. *ASME Journal of Biomechanical Engineering*, *130*(2), 021005.

Hutchison, K. A., Balota, D. A., Cortese, M.J., & Watson, J. M. (2008) Predicting semantic priming at the item level. *Quarterly Journal of Experimental Psychology*, *61*(7), 1036-66.

Idemaru, K., & Guion, S. (2008), Acoustic covariants of length contrast in Japanese stops. *Journal of Inter-national Phonetic Association*, *38*(2), 167-186.

Ijzerman, H., & Semin, G. R. (2009). The thermometer of social relations: mapping social proximity on temperature. *Psychological Science, 20*(10), 1214-1220.

Inagaki, T. K., & Eisenberger, N. I. (2013). Shared neural mechanisms underlying social warmth and physical warmth. *Psychological Science*, *24*(11), 2272-2280.

Iordanescu, L., Guzman, E., Grabowecky, M., & Suzuki, S. (2008). Characteristic sounds facilitate visual search. *Psychonomic Bulletin & Review*, *15*(3), 548-554.

Iordanescu, L., Grabowecky,M., Franconeri, S., Theeuwes, J., and Suzuki, S. (2010). Characteristic sounds make you look at target objects more quickly. *Attention, Perception, & Psychophysics*, *72*(7), 1736-1741.

Ito, T., Tiede, M., & Ostry, D. J. (2009). Somatosensory function in speech perception. *PNAS Proceedings of the National Academy of Sciences of the United States of America*, *106*(4), 1245-1248.

James, W. (1884). What is an Emotion? *Mind*, *9*(34), 188-205.

Jirak, D., Menz, M. M., Buccino, G., Borghi, A, M., Binkofski, F. (2010). Grasping language - a short story on embodiment. *Consciousness and Cognition*, *19*(3), 711-20.

Johnson, M. (1987). *The body in the mind: The bodily basis of meaning, imagination, and reason.* Chicago: University of Chicago Press.

Johnson-Laird, P. N. (1983). *Mental Models: Towards a Cognitive Science of Language, Inference, and Consciousness*. Cambridge: Cambridge University Press.

Jones, M., Kintsch, W., & Mewhort, J. K. M. (2006). High-dimensional semantic space accounts of priming. Journal of Memory and Language. *Journal of Memory and Language*, *55*(4), 534-552.

Jostmann, N. B., Lakens, D., & Schubert, T. W. (2009). Weight as an embodiment of importance. *Psychological Science*, *20*(9), 1169-1174.

Kaschak, M. P., & Glenberg, A. M. (2000). Constructing meaning: The role of affordances and grammatical constructions in sentence comprehension. *Journal of Memory and Language*, *43*(3), 508-529.

Kaschak, M. P., M. Aveyard, A. A. Blanchard, C. J. Madden, D. J. Therriault, R. H. Yaxley & R. Zwaan. (2005). Perception of motion affects language processing. *Cognition*, *94*(3), B79-B89.

Kaschak, M. P., Zwaan, R. A., Aveyard, M., & Yaxley, R. H. (2006). Perception of auditory motion affects language processing. *Cognitive Science*, *30*(4), 733-744.

Katz, D. (1989). The role if the temperature sensation in touch. In Krueger, L.E. (Ed.), *The World of Touch* (pp.165-177). Psychology Press.

Kawahara, S. (2015). The phonetics of sokuon, obstruent geminates. In Kubozono, H. (Ed.), *The Handbook of Japanese Language and Linguistics: Phonetics and Phonology* (pp.43-73). Mouton De Gruyter.

Kennett, S., Taylor-Clarke, M., & Haggard, P. (2001). Noninformative vision improves the spatial resolution of touch in humans, *Current Biology*, *11*(15), 1188-1191.

Kim, H.-G., Cheon, E.-J., Bai, D.-S., Lee, Y. H., & Koo, B.-H. (2018). Stress and Heart Rate Variability: A Meta-Analysis and Review of the Literature. *Psychiatry Investigation*, *15*(3), 235-245.

Kingston, J., Kawahara, S., Chambless, D., Mash, D., Brenner-Alsop, E. (2009). Contextual effects on the perception of duration. *Journal of Phonetics*, *37*(3), 297-320.

Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. New York: Cambridge University Press.

Koptjevskaja-Tamm, M. (Ed.). (2015). *The linguistics of temperature*. Amsterdam: John Benjamins Publishing Company.

Kouider, S., & Dupoux, E. (2001). A functional disconnection between spoken and visual word recognition: Evidence from unconscious priming. *Cognition*, *82*(1), B35-B49.

Krabs, R. U., Enk, R., Teich, N., & Koelsch, S. (2015). Autonomic effects of music in health and Crohn's disease: the impact of isochronicity, emotional valence, and tempo. *PLoS ONE*, *10*(5), e0126224.

Kurby, C. A., & Zacks, J. M. (2013). The activation of modality-specific representations during discourse processing. *Brain and Language*, *126*(3), 338-349.

Ladefoged, P., & Maddieson, I. (1996). *The Sounds of the World's Languages*. London: Blackwell Publishers.

Lakoff, G. (1987). *Woman, fire, and dangerous things: What categories reveal about the mind*. Chicago: University of Chicago Press.

Lakoff, G., & Johnson, M. (1980). The metaphorical structure of the human conceptual system. *Cognitive Science*, *4*(2), 195-208.

Lakoff, G., & Johnson, M. (1999). *Philosophy in the Flesh: The Embodied Mind and its Challenge to Western Thought*. New York, NY: Basic Books.

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), 211-240.

Landauer, T., McNamara, D. S., Dennis, S., and Kintsch, W (2006). Latent Semantic Analysis: A Road to Meaning. Mahwah, NJ: Erlbaum.

Lane, R. D., McRae, K., Reiman, E. M., Chen, K., Ahern, G. L., & Thayer, J. F. (2009). Neural correlates of heart rate variability during emotion. *NeuroImage*, *44*(1), 213-222.

Lang, A., Newhagen, J., & Reeves, B. (1996). Negative video as structure: Emotion, attention, capacity, and memory. *Journal of Broadcasting & Electronic Media*, *40*(4), 460-477.

Lederman, S.J., Klatzky, R.L. (2009). Haptic perception: a tutorial. *Attention, Perception, & Psychophysics*, *71*(7), 1439-1459.

Lee, Y. C, Lee, S. S., Zhang, W. C., Shih, Y. S., Yang, Y. H., Shen, H., Wu, C. C., Sune, S. T., Lee, B. J., & Huang, S. C. (2010). How much resting time is required before taking an ECG experiment? *Proceedings of the 2010 International Computer Symposium*, 846-849.

Lenggenhager, B, Azevedo, R.T., Mancini, A., & Aglioti, S.M. (2013). Listening to your heart and feeling yourself: effects of exposure to interoceptive signals during the ultimatum game. *Experimental Brain Research*, *230*(2), 233-41.

Levenson, R. W. (2014). The Autonomic Nervous System and Emotion. *Emotion Review*, *6*(2), 100-112.

Levinson, S. C., & Majid, A. (2014). Differential ineffability and the senses. *Mind & Language*, *29*(4), 407-427.

Liakakis, G., Nickel, J., & Seitz, R. J. (2011). Diversity of the inferior frontal gyrus–a meta-analysis of neuroimaging studies. *Behavioural Brain Research*, *225*(1), 341-347.

Liberman, A. M., & Mattingly, I. G. (1985). The motor theory of speech perception revised. *Cognition*, *21*(1), 1-36.

Liberman, A. M., Cooper, F. S., Shankweiler, D. P., & Studdert-Kennedy, M. (1967). Perception of the speech code. *Psychological Review*, *74*(6), 431-461.

Lindemann, O., Stenneken, P., Van Schie, H. T., & Bekkering, H. (2006). Semantic activation in action planning. *Journal of Experimental Psychology: Human Perception and Performance*, *32*(3), 633-643.

Louwerse, M. M., & Jeuniaux, P. (2010). The linguistic and embodied nature of conceptual processing. *Cognition*, *114*(1), 96-104.

Louwerse, M. M., Hutchinson, S., Tillman, R., & Recchia, G. (2015). “Effect size matters”: The role of language statistics and perceptual simulation in conceptual processing. *Language, Cognition and Neuroscience*, *30*(4), 430-447.

Lynott, D., Corker, K. S., Connell, L., & O’Brien, K. S. (2017). The effect of haptic and ambient temperature experience on prosocial behavior. *Archives of Scientific Psychology*, *5*(1), 10-18.

Mar, R. A. (2011). The neural bases of social cognition and story comprehension. *Annual Review of Psychology*, *62*,103-134.

Martin, A., Wiggs, C. L., Ungerleider, L. G., & Haxby, J. V. (1996). Neural correlates of category-specific knowledge. *Nature*, *379*(6566), 649-652.

Mastroberardino, S., Santangelo, V., & Macaluso, E. (2015). Crossmodal semantic congruence can affect visuo-spatial processing and activity of the fronto-parietal attention networks. *Frontiers in Integrative Neuroscience*, *9*, 45.

Mather, M., & Thayer, J. F. (2018). How heart rate variability affects emotion regulation brain networks. *Current Opinion in Behavioral Sciences*, *19*, 98-104.

Matsuda, Y., Kaneko, H., Taira, M. (2015). Relationship between the emotion and/or pupil fluctuation caused by the contents of Web and the impression of advertisement. *IEICE Technical Report*, *115*(36), 19-23.

McGurk H., & MacDonald J. (1976). Hearing lips and seeing voices. Nature, *264*, 746-748.

Meteyard, L., & Vigliocco, G. (2008). The role of sensory and motor information in semantic representation: A review. In Calvo, P., & Gomila, T. (Eds), *Handbook of Cognitive Science: An Embodied Approach* (pp.293-312). Elsevier.

Meteyard, L., Cuadrado, S. R., Bahrami, B., and Vigliocco, G. (2012). Coming of age: a review of embodiment and the neuroscience of semantics. *Cortex*, *48*(7), 788-804.

Meteyard, L., Zokaei, N., Bahrami, B., & Vigliocco, G. (2008). Visual motion interferes with lexical decision on motion words. *Current Biology*, *18*(17), R732-R733.

Miller, A. (2006). Watching viewers watch TV: Processing live, breaking, and emotional news in a naturalistic setting. *Journalism & Mass Communication Quarterly*, *83*(3), 511-529.

Möttönen, R., Watkins, K. E. (2009). Motor representations of articulators contribute to categorical perception of speech sounds. *Journal of Neuroscience*, *29*(31), 9819-9825.

Myung, J., Blumstein, S. E., & Sedivy, J. C. (2006). Playing on the typewriter, typing on the piano: Manipulation knowledge of objects. *Cognition*, *98*(3), 223-243.

Nabi, R. L., & Wirth, W. (2008). Exploring the role of emotion in media effects: An introduction to the special issue. *Media Psychology*, *11*, 1-6.

Nagano, A., & Shimada, M. (2014). Morphological theory and orthography: a representation of lexemes. *Journal of Linguistics*, *50*(2), 323-364.

Newell, A., & Simon, H. A. (1976). Computer science as empirical inquiry: symbols and search. *Communications of the ACM*, *19*(3), 113-126.

Newhagen, J. E., & Reeves, B. (1992). The evening's bad news: Effects of compelling negative television news images on memory. *Journal of Communication*, *42*(2), 25-41.

Nijhof, A. D., & Willems, R. M. (2015). Simulating Fiction: Individual Differences in Literature Comprehension Revealed with fMRI. *PLoS ONE*, *10*(2), e0116492.

Oba, R., Braun, A., Handke, J. (2009). The perception of Japanese geminates by native and non-native listeners. *Phonetician*, *92*, 9-29.

Oh, A., Duerden, E. G., & Pang, E.W. (2014). The role of the insula in speech and language processing. *Brain and Language*, *135*, 96-103.

Ooishi, Y., Mukai, H., Watanabe, K., Kawato, S., Kashino, M. (2017). Increase in salivary oxytocin and decrease in salivary cortisol after listening to relaxing slow-tempo and exciting fast-tempo music. *PLoS ONE*, *12*(12): e0189075.

Pecher D., Van Dantzig S., Boot I., Zanolie K., Huber D. E. (2010). Congruency between word position and meaning is caused by task-induced spatial attention. *Frontiers in Psychology*, *1*, 30.

Peltz, E., Seifert, F., DeCol, R., Dörfler, A., Schwab, S., & Maihöfner, C. (2011). Functional connectivity of the human insular cortex during noxious and innocuous thermal stimulation. *Neuroimage*, *54*(2),1324-1335.

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

Pulvermüller, F., & Hauk, O. (2006). Category-specific conceptual processing of color and form in left fronto-temporal cortex. *Cerebral Cortex*, *16*(8), 1193-1201.

Pulvermüller, F., Huss, M., Kherif, F., Moscoso del Prado Martin, F., Hauk, O., and Shtyrov, Y. (2006). Motor cortex maps articulatory features of speech sounds. *Proceedings of the National Academy of Sciences of the United States of America*, *103*, 7865-7870.

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), 498-508.

Pulvermüller, F., Shtyrov, Y., & Ilmoniemi, R. (2005). Brain signatures of meaning access in action word recognition. *Journal of Cognitive Neuroscience*, *17*(6), 884-892.

Pylyshyn, Z. W. (1973). What the mind's eye tells the mind's brain: A critique of mental imagery. *Psychological Bulletin*, *80*(1), 1-24.

Qian, W. (2016). Embodied cognition processing and representation of power words by second language learners with different proficiency levels. *Chinese Journal of Applied Linguistics*, *39(4)*, 484-494.

Radeau, M., Morais, J., & J. Segui. J. (1995). Phonological priming between monosyllabic spoken words. *Journal of Experimental Psychology: Human Perception and Performance*, *21*(6), 1297-1311.

Radvansky, G. A. & Zacks, J. M. (2014). *Event Cognition*. Oxford University Press.

Rall, J., & P. L. Harris. (2000). In Cinderella's slippers? Story comprehension from the protagonist's point of view. *Developmental Psychology*, *36*(2), 202-208.

Rizzolatti, G., & Craighero, L. (2004). The mirror neuron system. *Annual Review of Neuroscience*, *27*, 169-192.

Rizzolatti, G., Fogassi, L., & Gallese, V. (2001). Neurophysiological mechanisms underlying imitation and the understanding of action. *Nature Reviews Neuroscience*, *2*(9), 661-670.

Rosch, E., & Lloyd, B. B. (Eds.). (1978). *Cognition and categorization*. Lawrence Erlbaum.

Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology*, *8*(3), 382-439.

Russell, J. A. (2003). Core affect and the psychological construction of emotion. *Psychological Review*, *110*(1), 145–172.

Sadakata, M., Shingai, M., Brandmeyer, A., & Sekiyama, K. (2012). Perception of the moraic obstruent /Q/: a cross-linguistic study. *Proceedings of the Thirteenth Annual Conference of the International Speech Communication Association*, 891-894.

Sadakata, M., Shingai, M., Sulpizio, S., Brandmeyer, A., Sekiyama, K. (2014). Language specific listening of Japanese geminate consonants: a cross-linguistic study. *Frontiers in Psychology*, *11*(5), 1422.

Schachter, S., & Singer, J. (1962). Cognitive, social, and physiological determinants of emotional state. *Psychological Review*, *69*(5), 379–399.

Schomers, M.R., Kirilina, E., Weigand, A., Bajbouj, M., Pulvermüller, F. (2015). Causal Influence of Articulatory Motor Cortex on Comprehending Single Spoken Words: TMS Evidence, *Cerebral Cortex*, *25*(10), 3894-3902.

Šetić, M., & Domijan, D. (2007). The influence of vertical spatial orientation on property verification. *Language and Cognitive Processes*, *22*(2), 297-312.

Shaffer, F., & Ginsberg, J. P. (2017). An Overview of Heart Rate Variability Metrics and Norms. *Frontiers in Public Health*, *5*, 258.

Shaffer, F., McCraty. R., & Zerr, C. L. (2014). A healthy heart is not a metronome: an integrative review of the heart's anatomy and heart rate variability. *Frontiers in Psychology*, *5*, 1040.

Shigeno, S. (2017). Effects of Auditory and Visual Priming on the Identification of Spoken Words. *Perceptual and Motor Skills*, *124*(2), 549-563.

Simmons, W. K., Ramjee, V., Beauchamp, M. S., McRae, K., Martin, A., & Barsalou, L. W. (2007). A common neural substrate for perceiving and knowing about color. *Neuropsychologia*, *45*(12), 280202810.

Singer, T., Seymour, B., O'Doherty, J. P., Stephan, K. E., Dolan, R. J., & Frith, C. (2006). Empathic neural responses are modulated by the perceived fairness of others. *Nature*, *439*(7075), 466-469.

Slowiaczek, L. M., Nusbaum, H.C., Pisoni, D.B. (1987). Phonological Priming in Auditory Word Recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *13*(1), 64-75.

Sorabji, R. (1971). Aristotle on demarcating the five senses. *The Philosophical Review*, *80*(1), 55079.

Spielberger, C. D., Gorsuch, R. L., Lushene, R., Vagg, P. R., & Jacobs, G. A. (1983). *Manual for the State-Trait Anxiety Inventory*. Palo Alto: Consulting Psychologists Press.

Stanfield, R.A., Zwaan, R.A. (2001). The effect of implied orientation derived from verbal context on picture recognition. *Psychological Science*, *12*(2), 153–156.

Steinhubl, S. R., Wineinger, N. E., Patel, S., Boeldt, D. L., Mackellar, G., Porter, V., Redmond, J. T., Muse, E. D., Nicholson, L., Chopra, D., & Topol, E. J. (2015). Cardiovascular and nervous system changes during meditation. *Frontiers in human neuroscience*, *9*, 145.

Summerfield, A. Q. (1979). Use of visual information in phonetic perception. *Phonetica*, *36*(4-5), 314-331.

Tajadura-Jiménez, A., Väljamäe, A., Västfjäll, D. (2008). Self-representation in mediated environments: The experience of emotions modulated by auditory—vibrotactile heartbeat. *Cyberpsychology & Behavior*, *11*(1), 33-38.

Takahashi, W. (1998). Nihongo-no sokuon /Q/ nitsuite (Japanese moraic obstruent /Q/). *Journal of the Faculty of Education-Shinshu University*, *95*, 59-68.

Takakura, J., Nishimura, T., Choi, D., Egashira, Y., & Watanuki, S. (2015). Nonthermal sensory input and altered human thermoregulation: effects of visual information depicting hot or cold environments. *International Journal of Biometeorology*, *59*(10), 1453-1460.

Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology: Heart Rate Variability: Standards of Measurement, Physiological Interpretation, and Clinical Use, Circulation, Waltham, MA, USA, Mar. 1996, pp. 1043–1065.

Tettamanti, M., Buccino, G., Saccuman, M. C., Gallese, V., Danna, M., Scifo, P., Fazio, F., Rizzolatti, G., Cappa, S. F., Perani, D. (2005). Listening to action-related sentences activates fronto-parietal motor circuits. *Journal of Cognitive Neuroscience*, *17*(2), 273-281.

Therriault, D. J., Rinck, M., & Zwaan, R. A. (2006). Assessing the influence of dimensional focus during situation model construction. *Memory & Cognition*, *34*, 78-89.

Tiippana, K. What is the McGurk effect? (2014). *Frontiers in Psychology, 5*, 725.

Tucker, R., & Ellis, M. (2004). Action priming by briefly presented objects. *Acta Psychologica*, *116*(2), 185-203.

Unz, D., Schwab, F., & Winterhoff-Spurk, P. (2008). "TV news--The daily horror?": Emotional effects of violent television news. *Journal of Media Psychology: Theories, Methods, and Applications*, *20*(4), 141–155.

Valenza, G., Lanata, A., & Scilingo, E.P. (2012). The Role of Nonlinear Dynamics in Affective Valence and Arousal Recognition. *IEEE Transactions on Affective Computing*, *3*(2), 237-249.

Valins, S. (1966). Cognitive effects of false heart-rate feedback. *Journal of Personality and Social Psychology*, *4*(4), 400-408.

van Dijk, T. A., & Kintsch, W. (1983). Strategies of discourse comprehension. New York: Academic Press.

Vance, T. J. (1987). An Introduction to Japanese Phonology. Albany: State University of New York Press.

Vigliocco, G, Vinson, D. P., Lewis, W., & Garrett, M. F. (2004). Representing the meaning of object and action words: The featural and unitary semantic space hypothesis. *Cognitive Psychology*, *48*(4), 422-488.

Wang, H.M., & Huang, S. C. (2012). SDNN/RMSSD as a surrogate for LF/HF: A revised investigation. *Modelling and Simulation in Materials Science and Engineering*, 2012, 931943.

Watanabe, J., Kawaguchi, Y., K. Sakakura, K., & Ando, H. (2011). Heartbeat picnic: the workshop of feeling vibration. *Transactions of the Virtual Reality Society of Japan*, *16*(3), 303-306.

Watkins, K. E., Strafella, A.P., Paus, T. (2003). Seeing and hearing speech excites the motor system involved in speech production. *Neuropsychologia*, *41*(8), 989-994.

Weiskopf, D. A. (2010). Embodied cognition and linguistic comprehension. *Studies in History and Philosophy of Science Part A*, *41*(3), 294-304.

Williams, L. E., & Bargh, J. A. (2008). Experiencing physical warmth promotes interpersonal warmth. *Science*, *322*(5901), 606-607.

Wilson, M. (2002). Six views of embodied cognition. *Psychonomic Bulletin and Review*, *9*(4), 625-636.

Wilson, S. M., & Iacoboni, M. (2006). Neural responses to non-native phonemes varying in producibility: Evidence for the sensorimotor nature of speech perception. *Neuroimage*, *33*(1), 316-325.

Wilson, S. M., Saygin, A. P., Sereno, M. I., & Iacoboni, M. (2004). Listening to speech activates motor areas involved in speech production. *Nature Neuroscience*, *7*(7), 701-702.

Wilson-Mendenhall, C. D., Henriques, A., Barsalou, L. W., & Barrett, L. F. (2019). Primary interoceptive cortex activity during simulated experiences of thebody. Journal of *Cognitive Neuroscience*, *31*(2), 221-235.

Winter, B., & Bergen, B. (2012). Language comprehenders represent object distance both visually and auditorily. *Language and Cognition*, *4*(1), 1-16.

Winter, B., Perlman, M., & Majid, A. (2018). Vision dominates in perceptual language: English sensory vocabulary is optimized for usage. *Cognition*, *179*, 213-220.

Wortman, J., Donnellan, M. B., & Lucas, R. E. (2014). Can physical warmth (or coldness) predict trait loneliness? A replication of Bargh and Shalev (2012). *Archives of Scientific Psychology*, *2*(1), 13-19.

Xu, J., Kemeny, S., Park, G., Frattali, C., & Braun, A. (2005). Language in context: emergent features of word, sentence, and narrative comprehension. *NeuroImage, 25*(3), 1002–1015.

Xue, J., Marmolejo-Ramos, F., and 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.

Yoshida, N. (2015). *Tear drop glasses*. Linz: Ars Electronica.

Zhong, C. B., & Liljenquist, K. (2006). Washing away your sins: threatened morality and physical cleansing. *Science*, *313*(5792), 1451-1452.

Zhou, Y., & Watanabe, J. (2020). Non-informative Visual Resolution of the Body Modulates Tactile Acuity. *Transactions of the Virtual Reality Society of Japan*, *25*(4).

Zwaan, R. A. (2004). The Immersed Experiencer: Toward an Embodied Theory of Language Comprehension. In B. H. Ross (Ed.), *The psychology of learning and motivation: Advances in research and theory* (pp. 35-62). Elsevier.

Zwaan, R. A., & Kaschak, M. P. (2008). Language in the brain, body, and world. In Robbins P., & Aydede, M. (Eds.), *Cambridge Handbook of Situated Cognition* (pp. 361-381). Cambridge University Press.

Zwaan, R. A., & Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychological Bulletin*, *123*(2), 162-185.

Zwaan, R. A., & Taylor, L. J. (2006). Seeing, acting, understanding: Motor resonance in language comprehension. *Journal of Experimental Psychology: General*, *135*(1), 1–11.

Zwaan, R. A., Langston, M. C., & Graesser, A. C. (1995). The construction pf situation models in narrative comprehension: an event indexing model. Psychological Science, *6*(5), 292-297.

Zwaan, R. A., Magliano, J. P., & Graesser, A. C. (1995). Dimensions of situation-model construction in narrative comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*(2), 386–397.

Zwaan, R. A., Radvansky, G. A., Hilliard, A. E., & Curiel, J. M. (1998). Constructing multidimensional situation models during reading. *Scientific Studies of Reading*, *2*(3), 199-220.

Zwaan, R. A., Stanfield, R. A., & Yaxley, R. H. (2002). Language comprehends mentally represent the shapes of objects. *Psychological Science*, *13*(2), 168-71.

Zwaan, R., & Ross, B.H. (2004). The immersed experiencer: Toward an embodied theory of language comprehension. In Medin, D. L. (Ed.), *The Psychology of Learning and Motivation* (pp.35-62). San Diego: Academic Press.

It is noted that this thesis interpolates materials from the following published first-author papers [1-4]. Partial results of the presented work have been published in [1, 3, 4].

[1] Zhou, Y., Nakamura, Y., Mugitani, R., & Watanabe, J. (2021). Influence of prior auditory and visual information on speech perception: Evidence from Japanese singleton and geminate words. *Acoustical Science and Technology, 42*(1), 36-45. doi: 10.1250/ast.42.36

[2] Zhou, Y., & Watanabe, J. (2020). Non-informative visual resolution of the body modulates tactile acuity. *Transactions of the Virtual Reality Society of Japan, 25*(4), 315-318. doi: 10.18974/tvrsj.25.4_315

[3] Zhou, Y., Murata, A., & Watanabe, J. (2020). The calming effect of heartbeat vibration. *Proceedings of the IEEE Haptics Symposium (HAPTICS)*, 677-683. doi: 10.1109/HAPTICS45997.2020.ras.HAP20.157.5a2e1551

[4] Zhou, Y., Ho, H-N., & Watanabe, J. (2017). Perceptual-semantic congruency facilitates semantic discrimination of thermal qualities. *Frontiers in Psychology, 8*, 2113. doi: 10.3389/fpsyg.2017.02113