Sound Symbolic Words in Japanese Talk-in-interaction

By

Yumiko Shibata

A dissertation submitted in partial fulfillment of

the requirements for the degree of

Doctor of Philosophy

(Japanese)

at the

UNIVERSITY OF WISCONSIN-MADISON

2017

Date of final oral examination: 3/31/2017

The dissertation is approved by the following members of the Final Oral Committee:

Junko Mori, Professor, Japanese
Naomi McGloin, Professor, Japanese
Naomi Geyer, Associate Professor, Japanese
Douglas W. Maynard, Professor, Sociology
Cecilia E. Ford, Honorary Associate/Fellow Emeritus, Sociology and English

Acknowledgments

This dissertation could not have been completed without the support and assistance I have received from many individuals. In the first place, I would like to express my deepest appreciation to my advisor, Professor Junko Mori, for her professional guidance over the last seven years. She has read each chapter again and again, provided me with valuable insight into various aspects of my work, and supported me in shaping my ideas for this dissertation. Without her guidance and persistent help, this dissertation would not have been possible. My appreciation also goes to other committee members. Professor Naomi McGloin, Professor Naomi Geyer, Professor Doug Maynard, and Professor Cecilia Ford have always shown sincere interest in my data and provided me with insightful feedback, which edified my scholarship.

I am also grateful to have had the opportunity to meet the following scholars and to talk with them about my study: Dr. Charles Goodwin, Dr. Makoto Hayashi, Dr. Mark Dingemanse, Dr. Noriko Iwasaki, Dr. Hiromichi Hosoma, Dr. Shoichi Iwasaki, Dr. Ikuyo Morimoto, and Dr. Hiroki Maeda. They have given me thought-provoking comments at different stages of my work. I'd particularly like to thank Dr. Mark Dingemanse for giving life to my dissertation through his support of my ideas. I am also thankful to Manuel Herrero-Puertas, Gary Koutsoubis, and Mike Haen, who proofread my chapters thoroughly. My appreciation also goes to the anonymous participants of this study. This study would not have happened without their generous participation.

My thanks also goes out to Dr. Takako Nakakubo for her advice and encouragement with regard to my Ph.D. student life. I am also thankful to all of those with whom I have had the pleasure of working with during my graduate study in Madison: Chiharu Shima, Jae Takeuchi,

Jun Xu, Hironori Nishi, Yuki Arita, and Akiko Imamura. I would especially like to thank Yuki and Akiko, who have spent countless hours with me in the library sharing joy and pain and laughter in addition to lots and lots of ideas. Let me also express my special thanks to Peter Miliczky, my violin teacher in Madison, for bringing a breath of fresh air into my life.

I would also like to express my heartfelt gratitude to my professors at the University of Auckland in New Zealand. My former advisor Professor Fay Wouk introduced me to the fun of analyzing everyday conversation data. Professor Frank Lichtenberk is the one who made me realize that Japanese SSWs can be a subject of research. Without taking his typological-functional linguistics course, this dissertation would not be what it is today. I also would like to thank my senior colleagues at the University of Auckland, Hideyuki Sugiura and Natsuko Shintani, for their advice and encouragement. Their sincere attitudes towards research have commanded my respect.

My deep gratitude also goes to my professors at Dokkyo University in Japan. Professor Yuka Iijima opened a window into academia in English for me, and Professor Yoshiro Asayama first taught me the delight which could be found by discovering rules in languages.

I would also like to express my appreciation to the members of the Japanese program at the University of California at Davis for their understanding of my situation and encouragement. I am also very thankful to my friends, especially Sachi Takahashi and Megumi Kuromiya, for our long lasting friendship. I also cannot forget to mention my late friend Atsuko Edakubo, who taught me to live my life in my own way.

Finally, I truly appreciate the unconditional support that my parents provided me during this journey. I am so thankful for their faith in me as I was searching for my path.

Abstract

This dissertation examines deployment of sound symbolic words (SSWs), also known as mimetics and ideophones, in Japanese face-to-face interaction. By analyzing data from folktales, narratives, and written texts, for example, a large number of previous studies investigated SSWs across languages, and attempted to pin down the meaning of each instance of an SSW. One intriguing fact is that they have identified intrinsic semantic characteristics of SSWs which are contrasting, such as "vivid" (e.g., Doke, 1935; Kita, 1997; Schourup, 1993) and "specific" (e.g., Childs, 1994) as opposed to "ambiguous" (Noss, 2003) and "elusive" (Bartens, 2000). Further, scholars found that SSWs are "quintessentially social" (Childs, 1994:63) and "the closest linguistic substitute for a non-verbal, physical act" (Kunene, 2001:183), emphasizing the importance of taking the context of use into consideration, yet no studies have examined SSWs in face-to-face interaction, except for Dingemanse (2011, 2013, 2014).

Through close examination of naturally occurring conversation, this study sheds light on how deployment of SSWs operates in social interaction. Primarily informed by gesture researchers such as Adam Kendon and David McNeill who described speakers' bodily conduct co-occurring during their spontaneous speech, the study first examines how turns that incorporate SSWs are constructed, focusing on occurrences of various multimodal resources, such as hand gestures and vocal gestures. Building on this analytical foundation, it further examines what speakers can achieve through the deployment of SSWs in interaction. More specifically, one of the analytical chapters demonstrates how SSWs work as a resource to depict and concretize objects that do not exist in the interactional space. The chapter then demonstrates that SSWs can serve as a provisional element in word search activities. Another analytical

chapter reveals how SSWs serve as a resource to render speakers' emotional stances which had not been clearly expressed by the use of linguistic resources other than SSWs. Moment-by-moment analyses further show how the contrastive semantic characteristics introduced above can evolve over the course of the interaction. This study highlights the fluid and dynamic nature of SSWs, and provides empirical analyses of actual uses of SSWs, which are believed to be found in all world languages.

Table of Contents

Acknowledgements i
Abstract iii
Table of Contents v
List of Tables vii
Transcription Conventions viii

Part 1 Introduction 1

Chapter 1 Introduction 2
1.1 Objectives and purpose of this study 2
1.2 Terminology 6
1.3 Previous studies 8
1.4 Data and methodology 14
1.4.1 Data 14
1.4.2 Transcription and English translation 17
1.4.3 Methodology 18
1.5 Organization of the dissertation 20

Part 2 Preliminary to Analysis:
How Depictions through Sound Symbolic Words Are Embodied 23

Chapter 2 Sound Symbolic Words and Bodily Conduct 24
2.1 Introduction 24
2.2 Gesture studies from cognitive and interactional perspectives 27
2.2.1 Cognitive perspectives 27
2.2.2 Interactional perspectives 29
2.3 Creative uses of conventional SSWs and invented SSWs 31
2.3.1 With hand gestures and distinctive prosodic design 31
2.3.2 Without hand gestures nor distinctive prosodic design 40
2.4 Notes on facial gestures 43
2.5 Summary of chapter 2 44

Chapter 3 Sound Symbolic Words as Vocal Gestures 45
3.1 Introduction 45
3.2 Sound symbolism and iconicity 49
3.3 Meanings of the SSW *pon* in previous studies 51
3.4 Motions 54
3.5 Invisible concepts 64
3.6 Motionless objects 73
3.7 Summary of chapter 3 80

Part 3 Analysis:
What Speakers Can Do through Deployment of Sound Symbolic Words in Talk-in-interaction **82**

Chapter 4 Concretizing Absent Objects through Deployment of Sound Symbolic Words **83**
4.1 Introduction 83
4.2 Instant depiction of absent objects 85
4.3 Possibilities and limitations of SSWs 92
4.4 SSWs as a provisional element in word search activities 106
4.4.1 Outlining word search activities 107
4.4.2 How SSWs operate as a provisional element in word search activities 110
4.5 Summary of chapter 4 124

Chapter 5 Rendering Emotional Stances into Sound Symbolic Words **125**
5.1 Introduction 125
5.2 SSWs as a readily available resource for rendering emotional stance 129
5.3 SSWs as a resource for the recipient to display understanding 151
5.4 Summary of chapter 5 164

Part 4 Conclusion **166**

Chapter 6 Conclusion **167**
6.1 Contributions of this study 167
6.2 Implications and possible future directions 171

References **175**

List of Tables

Table 1.1 Description of Each Conversation 15
Table 2.1 The Numbers Distributions of SSWs That Co-occur With and Without Gestures 24

Transcription Conventions

1. Transcription symbols

[	starting point of overlapped talk
]	ending point of overlapped talk
=	latching
(.)	micro pause
(0.0)	length of silence
↑	high in pitch
↓	low in pitch
> <	increase in tempo
< >	decrease in tempo
___	emphasis
CAPS	louder volume than surrounding talk
:	lengthened sound
.	falling intonation
,	continuing intonation
?	rising intonation
¿	rising intonation but not as high as "?"
-	cut-off
( )	unintelligible stretch
(word)	transcriber's uncertain hearings
(( ))	commentary by the transcriber
h	exhalation
.h	inhalation
(h)	exhalation, laughter, within a word
→	turns including sound symbolic words

2. Abbreviations

CP	copula
FP	final particle
LK	nominal linking particle
N	nominalizer
NG	negation
OB	object marker
PT	past tense
Q	question marker

QT	quotative particle
SB	subject marker
SSW	sound symbolic word
TAG	tag-like expression
TP	topic marker

Part 1

Introduction

Chapter 1

Introduction

1.1 Objectives and purpose of this study

The Japanese language has a word class called sound symbolic words (SSWs), also known as mimetics and ideophones, that consists of phonomimes (*giongo* 'sound mimicking words' and *giseigo* 'voice mimicking words'), phenomimes (*gitaigo* 'state mimicking words') and psychomimes (*gijoogo* 'psychological state mimicking words'). As the names indicate, SSWs mimic natural sounds or depict manners and states of the referent in sound. This study examines speakers' deployment of SSWs while coordinating visible hand gestures and audible sound qualities in face-to-face interaction. Let us present a brief segment below in which the speaker Mie suggests an imaginary function of a pointer stick through the SSW ↑>*hyun*<.

(1.1) (elicited from (4.2) in Chapter 4)

Figure 1.1

Figure 1.2

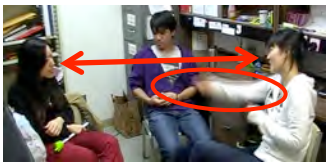


Figure 1.3

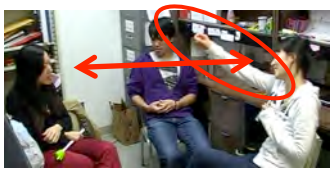


Figure 1.4

63 Mie: → *botan toka oshite ↑>hyun< te nobiru [yatsu toka*
button etc. push.and SSW QT extend one etc.
'how about if it extends like ↑>*hyun*< by pushing something like a button'

The deployment of the SSW >*hyun*< concretizes the object that does not exist in the interactional space. Specifically, vocal quality is differentiated from the surrounding talk by accelerating the word in a high pitch, thereby depicting the object. I call the voice quality changes "vocal gestures" (Voeltz & Kilian-Hatz, 2001a:3). The speaker further makes use of visible body movements and gaze. SSWs depict a wide range of attributes of invisible objects as various multimodal resources are simultaneously deployed.

SSWs are also capable of depicting the emotional stances of participants. The next segment illustrates that the speaker designs the turn wherein she visibly depicts her emotional reaction of frustration using her hands shaking around her head. Further, the vocalization of the SSW is emphasized by stressing the word in a louder volume.

(1.2) (elicited from (5.6) in Chapter 5)

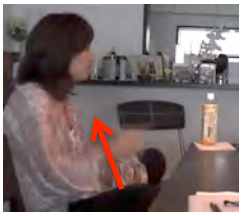


Figure 1.5

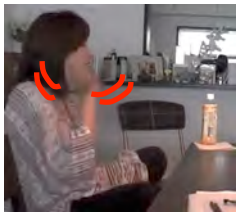


Figure 1.6

50 Chie: → (.)[>*UUUn*< *tte nan* [*noni*
SSW QT become even.though
'I become like >*UUUn*< and further'

What has to be highlighted here is that, in face-to-face interaction, speakers design the sound qualities and visible gestures in accordance with how the turn unfolds and what is being depicted.

SSWs are not only limited to Japanese but found across languages. The descriptive function of adverbs led to the recognition of SSWs being associated with vividness. In studying Bantu languages, Doke (1935) defined ideophones as follows:

> A vivid representation of an idea in sound. A word, often onomatopoeic, which describes a predicate, qualificative or adverb in respect to manner, colour, sound, smell, action, state or intensity (p. 118).

Doke's definition is highly influential across SSWs studies; the word "vivid" has permeated throughout the description of SSWs in previous studies (e.g., de Jong, 2001; Fortune, 1963; Iwasaki, Brown, Kita & Vinson, 2013; Johnson, 1976; Kita, 1997; Klamer, 2001; McGreger, 2001; Okpewho, 1992; Schourup, 1993; Tamori, 2002). Several related adjectives have also been adopted in describing semantic characteristics of SSWs. For example, scholars described the semantics of SSWs as "eloquent" (Johnson, 1976), "specific" (Amha, 2001; Childs, 1994; de Jong, 2001; Watson, 2001), and "dramaturgic" (Voeltz & Kilian-Hatz, 2001a). In a related move, scholars in Japanese linguistics identified characteristics of SSWs in discourse. Schourup (1993) described Japanese SSWs by using the following terms: "ostentatiousness" (p. 95), "condensation" (p. 96), and "concreteness" (p. 96). Schourup argued that use of SSWs made the meaning of the discourse condensed and concrete, thereby helping the recipient to invoke the detailed situation.

However, some other scholars characterized SSWs with adjectives that contrast with the above mentioned, such as "ambiguous" (Noss, 2003) and "elusive" (Bartens, 2000; Noss, 2003). Such contrasting characteristics of SSWs were identified while examining data, primarily from folktales, narratives, and written texts, for example, and describing the lexical meaning of each SSW. In other words, scholars found those semantic characteristics as intrinsic to SSWs, without fully addressing how these characteristics arise while interaction unfolds. In fact, some scholars noticed that SSWs are "quintessentially social" (Childs, 1994:63) and "the closest linguistic substitute for a non-verbal, physical act" (Kunene, 2001:183). Furthermore, Bartens (2000) pointed out that "the semantics of ideophones can be quite elusive and become clear only in the

particular context in which they are used" (p. 26). Yet, no scholar has paid attention to actual use of SSWs embedded in a social interactional context, except for Dingemanse (2011, 2013, 2014).

This dissertation fills this gap and sheds light on how deployment of SSWs operates in naturally occurring interaction. As Barten's statement about the significance of the context affirms, SSWs never appear without having a context in social interaction. Social interaction is on-going processes through which participants build their own turns while monitoring how others participate in interaction.

The purpose of this study is twofold. This study first aims to reveal how SSWs operate as depictive resources as speakers coordinate hand gestures and vocal gestural aspects of SSWs (Chapters 2 and 3). Previously, SSWs were confirmed to co-occur with hand gestures (e.g., Kita, 1997) and were characterized as vocal gestures. The remaining issue that needs to be addressed is how these multimodal characteristics operate upon deployment of SSWs in interaction. Once this issue is clarified, this study further aims to demonstrate how participants in talk-in-interaction accomplish interactional goals through deployment of SSWs (Chapters 4 and 5). As noted, scholars attempted to determine the meaning of each individual SSW, up until recently. However, given that language is a primary vehicle of interaction in social context, gaining understanding of SSWs in social interaction is a crucial step toward deeper insight into SSWs.

This study further elucidates the seemingly contrasting characterizations of SSWs pointed out above, by demonstrating that SSWs can be either specific or ambiguous depending on which point of unfolding interaction the analyst looks at. Interaction is a process that develops over time. Thus, deployment of SSWs in certain sequential contexts does not necessarily make understanding clear for the recipient. The participants may go through steps in which they

mutually make sure they understand what has been depicted through an SSW. In this manner, this study provides a thorough analysis of SSWs in naturally occurring face-to-face interaction.

1.2 Terminology

Before summarizing key studies that have impacted our understandings of SSWs, I provide a summary of related terminologies that have been used in previous studies. The Japanese terminology *onomatope* 'onomatopoeia' is frequently used among Japanese scholars, but what it refers to is different from the original meaning. The translated *onomatope* is often used to cover phonomimes (*giongo* 'sound mimicking words' and *giseigo* 'voice mimicking words') and phenomimes (*gitaigo* 'state mimicking words') (e.g., Tamori & Schourup, 1999; Ono, 2007), while the original terminology onomatopoeias only refer to words that depict natural sounds.

Some studies (e.g., Akita, 2012; Hasada, 2001; Hayashi, 2003b, 2005b; Kita, 1993, 1997, 2001; Toratani, 2005, 2006, 2009) employed the terminology mimetics or mimetic words, whose focal point is that the word mimics or imitates the state of objects or manner of actions. Some other studies employed the terminology ideophones whose main concern is that speakers represent their idea in sound (e.g., Childs, 1994; Dingemanse, 2011; Voeltz & Kilian-Hatz, 2001a).

While there are various terminologies used in the studies of Japanese SSWs, Hamano's (1998) statement prompted me to choose the name sound symbolic words over other terminologies. Hamano pointed out that a large number of "mimetic words are sound symbolic rather than onomatopoeic… Only a fraction of them imitate sounds" (p. 2). When a word is onomatopoeic, the word imitates actual sounds that you hear. In contrast, when a word is sound symbolic, the word symbolizes or signifies the referent of the word, which is not associated with

natural sound. Hamano's statement does not correspond with Doke's definition introduced earlier, but instead corresponds with the instances of SSWs in my database, most of which are non-onomatopoeic. A segment from my database is shown below. The SSW *be* (line 5) is not based on natural sound but rather is sound symbolic. Prior to the segment, Kimi explained that her house was adjacent to her grandmother's house.

(1.3) [OKS.2418]

1 Kimi: *nanka uchi, (e)tto ie ga atte:,*
somehow house, well house SB exist.and
'well, my house, there is the house, and'

2 *watarirooka?*
connecting.corridor
'a connecting corridor?'

3 Yuko: [*un*
'uh huh'

4 Kimi: [*mitaina,*
like
'something like'

5 → *nanka ki no ita ga koo be tte atte*
like wood LK board SB like.this SSW QT exist.and
'like, a wood board is lying like *be*, and'

In this segment, Kimi explains that there is a wood board like a connecting corridor, while the state of the wood placed there is depicted with the SSW *be.* Kimi's utterance does not mean that the wood board lying between two houses makes the sound of *be*. A large number of instances in my database are of this non-onomatopoeic type.

This dissertation classifies SSWs into three categories, conventional SSWs, creative SSWs, and invented SSWs, to organize the variety of SSWs and show how frequently speakers deploy unique SSWs that are not listed in dictionaries. First, *conventional SSWs* refer to the

SSWs which are found in specialized dictionaries of SSWs. Among more than ten available dictionaries, I have selected the following dictionaries that contain abundant items and accompanying examples, in order to see if SSWs in question are considered conventional SSWs. One is Kakehi, Tamori, and Schourup (1996a, 1996b), two volume dictionary written in English, containing approximately 3700 entries and examples for each SSW. The other is Ono (2007), published approximately a decade later, written in Japanese, that covers 4700 words including dialects from modern as well as classic uses of SSWs. I have encountered some words that are listed in Ono (2007) but not listed in Kakehi, et al. (1996a, 1996b). The difference in the covered items most likely comes from the differences in the data source and/or the differences seen across years of publication. Second, *creative use of conventional SSWs*, or simply *creative SSWs*, refer to conventional SSWs produced in distinct phonetic design, for example, by prolonging the vowel or by changing the pitch of the part of or the whole words. For example, ↑*po:n* and <*ponponpon*> (see Chapter 3) are creative variations of the conventional SSW *pon*. Finally, *invented SSWs* are SSWs that are idiosyncratic SSWs that cannot be found in any dictionaries. They are original and unique strings of sounds.

1.3 Previous studies

Scholars have investigated semantic characteristics of SSWs taking into consideration the contexts – either linguistic or social– in which SSWs are placed, mainly using data other than naturally occurring conversation. Meanwhile, the authors of dictionaries have pointed out atypical uses of SSWs. Hida and Asada (2002) noted that "the authors of SSWs could freely create their own versions of SSW, not following successions of characters that deviate from the typical patterns. In that way, the author expresses their feelings and adds some nuances to the

word" (p. v).[1] Moreover, Ono (2007) provided his observation "when the speaker/writer makes up SSWs, the meaning may be unclear to the hearers."[2] Ono further continued that "there are many cases in which the meaning of SSWs can be difficult to understand for someone who was not present in the situation where the SSW was produced" (p. 22).[3] These statements clearly indicate the necessity of examining SSWs in actual interaction in order to determine how the meaning of SSWs can be difficult to understand, and how participants deal with those atypical uses of SSWs when participating in talk-in-interaction.

Hasada (2001) raised questions about dictionary definitions and stressed the importance of considering the context. Hasada's study shares the issue in question with the current study, while data and methodology is different. Hasada (2001) conducted the analysis of the two different pychomimes—*dokiQ*[4] and *gyoQ*—both of which are defined as the feeling of *being afraid* and *being in fear*. Hasada pointed out that they are defined in the same way, although they are not interchangeable in sentences. Aiming to differentiate them from one another, Hasada defined the meaning of several psychomimes, including *dokiQ* and *gyoQ*. The data were taken from various sources, such as newspapers, TV commercials, popular songs, among others. For each SSW, Hasada analyzed approximately 10 pieces of data and formulated a definition. An example is shown in (1.4) below, which is her definition of the SSW *gyoQ.* Hasada's definition concerns the context of its use and the speaker's feeling upon using it. Noticeably, however,

---

[1] The original Japanese wording is*: sakka ga jibun dake no kansei ni yotte tsuujoo dewa nai mojiretsu o mochiite dokuji no imi/nyuansu o motase tari, hanashite ga kanjoo ni makasete oto o jiyuu ni soosaku shitari shite hyoogen surukoto wa juubun ni ariuru koto de aru.*

[2] The original Japanese wording is*: jibun de katte ni kanyoo to kotonaru onomatope o amidasu to tsutawari nikukattari.*

[3] The original Japanese wording is*: sonoba ni iawasenai hito ga ato kara kiite mo wakaranaka ttari toiu baai ga tata ariuru.*

[4] Q represents the syllable final moraic obstruent, or double consonants. It is also called *sokuon* in Japanese.

some words found in the definition are vague, such as "bad" and "a very short time," which would differ across individuals.

(1.4) *GyoQ*

(a) X feels something for a very short time
(b) because X thinks something
(c) sometimes a person thinks something like this:
(d) "something happens now
(e) I know: this is bad
(f) I did not know before that something like this could happen
(g) I can't think now
(h) I can't do anything now because of this"
(i) because of this, this person feels something bad for a very short time
(j) X feels something like this
(k) because X thinks something like this

(Hasada, 2001:228-29)

Furthermore, the expression "feels something like this" in (j) evidently shows the difficulty of pinning down the meaning of *gyoQ*. Coincidentally, a data segment analyzed in Chapter 5 includes the SSW *gyo*. My analyses will show that participants gradually build their understanding of *gyo* through the ongoing interaction rather than through pre-existent knowledge of the word.

One of the most recent studies on Japanese SSWs approached the semantics issue by examining collocational patterns associated with mimetics. Akita (2012) provided a Frame Semantic analysis in which "the meanings of linguistic expressions are understood in terms of the whole background situation or knowledge (or 'frame') to which they pertain" (p. 69). Significantly, Akita's study that analyzed written data[5] raises a question with respect to how recipients shape their understanding of SSWs. The following set of examples in (1.5) reveals that linguistic analysis using Frame Semantics appears to treat language as final products, not as an

[5] Akita's analysis is based on two databases, containing transcribed written and spoken language, which are linked to software for morphological analysis, called *Chakoshi* and *Chasen*.

ongoing process. Akita argues that a mimetic can evoke more than one frame, with the help of the preceding noun and the following verb. The squared parts in (1.5) come into questions.

(1.5)[6] a. *soko-ni-wa nogiku-ya-kikyoo-ga saki-midarete,*
there-to-TP wild.chrysanthemum-and-bellflower-SB bloom.and
aki-no tyoo-ga *hirahira-to mat-te-ita.*
fall-LK butterfly-SB MIM-QT dance.and-was
'In the place, wild chrysanthemums and bellflowers were blooming all over, and autumnal butterflies were dancing flutteringly.'
(Kido Okamoto, Tamamo-no-mae [Tamamo-no-Mae])
(Akita, 2012:79, gloss is simplified)

b. *soko-e hontooni kaze-to-tomoni itiyoo-no-tegami-ga kare-no-temoto-e*
there-to really wind-with-together one-LK-letter-SB he-LK-at.hand-to
hirahira tonde-kita.
MIM fly.and-came
'To the place, a letter came flying flutteringly to [his] hands literally in the wing.'
(Osamu Dazai, Sarumen-kanzya [Monkey-faced youth])
(Akita, 2012:79, gloss is simplified)

The mimetics in (1.5a) and (1.5b) function as adverbs that modify the subsequent verbal predicates, *mat-* 'to dance' in (1.5a) and *tob-* 'to came flying' in (1.5b). According to Akita, the verb *mat-* 'to dance' in (1.5a) and the verb *tob-* 'to fly' in (1.5b) evoke the frame of 'motion.' That is, in the above examples, the evoking verbs follow the mimetic. The Frame Semantic analysis appears not to take the positional relationship between the mimetic and the collocated words into consideration, and does not consider the process of how the reader interprets the text. Written language remains the same across time and space (Olson & Torrance, 1983). Moreover, reading texts is "a slow, deliberate, editable process" (Chafe, 1985:105). Thus, when one reads the sentences in (1.5), he/she follows each word from the beginning to the end one-by-one. It will turn out that the reader's interpreting process of the mimetic *hirahira* has to be

[6] The following is the list of abbreviations used in the word-by-word gloss: LK=linking nominal, MIM=mimetics, QT=quotative particle, SB=subject marker, TP=topic marker.

undetermined until the verb, which evokes the frame of 'motion,' comes into his/her eyes. Thus, Akita's study raises the following question: If the verb placed at the end of the sentence evokes the frame, what does the prior SSW do in the sentence? This dissertation poses this question and analyzes face-to-face ongoing interaction, paying primary attention to how each turn unfolds.

In exploring how language production is linked with human cognition, Kita (1997) studied mimetics by categorizing them into nominal mimetics and adverbial mimetics, using invented sentential examples. Kita found that nominal mimetics represent both the analytic and the affecto-imagistic dimensions of meaning, while adverbial mimetics only represent the affecto-imagistic dimension. According to Kita, the analytic dimension is associated with "ordinary semantics" (p. 379), conveying descriptive information "explicitly asserted or denied and, in the most favourable instances at least, it can be objectively verified" (Lyons, 1977:52, cited in Kita, 1997:386). In contrast, the affecto-imagistic dimension activates "the affective, emotive, and perceptual" (p. 387) aspects of experience and makes "the subjective effect of evoking an image or 're-experiencing'" (p. 387). By analyzing spoken data taken by a re-telling task, Kita further found that Japanese adverbial mimetics are "tightly coupled with iconic gestures" (Kita, 1997:359), explaining that "mimetics are almost always accompanied by a stroke … of a coexpressive gesture" (p. 392). A stroke is one of the gesture phases with which the central meaning is expressed (McNeill, 1992). Put differently, visible multimodal resources play prominent roles in deployment of SSWs.

Vocalization is also inseparable from actual deployment of SSWs. The human voice has "the wide range of precisely controllable variability it puts at its owner's disposal" (Burrows, 1990:29). The vocalization of SSWs creates moments where such disposals come into play, and thus deserves to be called vocal gestures. Audible human voice has been studied under the notion

of iconicity (Perlman, 2010; Shintel et al., 2006). For example, Perlman (2010) found that "descriptions of fast events were spoken at a faster rate than descriptions of slow events" (p. 253). In actual interaction, such descriptive vocalization is fundamental to the deployment of SSWs.

In analyzing data from naturally occurring conversation in Siwu, a language spoken in part of Ghana, Dingemanse (2014) centered on newly created ideophones that perform depiction in conversation. Based on the data analysis, he claimed that "there is a transparent link to the affordances of speech as they are exploited in everyday language use" (p. 401), and suggested the necessity of conducting thorough examinations of SSWs in everyday conversation. Furthermore, along with analyses of newly created ideophones in interaction, he presented a segment in which he interviewed a participant of the interaction. In responding to the question as to how native speakers understand creative ideophones, the participant explained that they took the meaning "[f]rom the conversation" (p. 397). My study demonstrates how native speakers of Japanese handle the meaning of SSWs through ongoing everyday face-to-face interaction where all the available multimodal resources (i.e., hand, body, and voice) operate along with the linguistic resource of SSWs.

In summary, scholars have confirmed that SSWs are in some way related to depiction, however, there remain unsolved issues. First, how do hand gestures and vocal gestural aspects of SSWs simultaneously operate in face-to-face interaction? Second, by making use of vocal gestural aspects of SSWs that accompany hand gestures, what can speakers accomplish in face-to-face interaction? Through the examination of naturally occurring conversation data, this study clarifies these issues.

1.4 Data and methodology

1.4.1 Data

The database of this study consists of 31 sets of face-to-face video-recorded ordinary conversation among native speakers of Japanese, approximately 30 hours in total. Supplementary, the interaction was also audio-recorded to capture the detailed vocal presentations. Each set of conversation is conducted between two to four friends or colleagues. In total, 70 native speakers of Japanese, 17 male and 53 females, in their 20's to 50's, participated in this research. This database was collaboratively constructed with co-researchers who investigated how social interaction is organized using verbal and non-verbal resources in Japanese conversation. The participants were recruited by word of mouth; the researchers contacted friends and acquaintances in the US and in Japan, and asked if they could bring friends and participate data recording sessions for this research. The researchers also asked those friends to ask their friends to participate in the research. Once they showed interest, the participants and the researchers have arranged on a time and place to meet for data recording.

At least one of the researchers visited the site of the data recording, and set up the devices: a camcorder and an IC recorder. When the participants arrived, researchers asked the participants to have free conversation; no conversation topics were provided. Before the data recording began, researcher(s) handed the participants the consent form, explained (1) that researchers examined face-to-face interaction without telling them the specific objectives of the study, (2) how the data are stored, and (3) their right to withdraw from this research at any time. Participants then put their signature on the consent form to show their acknowledgement. The researchers then started the recording. The researchers agreed with the participants regarding a

time to finish recording, left the room, and came back to the room at the arranged time. When the researchers came back to the room, the recording was completed.

Among the 31 sets of conversation data, some data sets contained more occurrences of SSWs than others. Put differently, some participants use SSWs quite often, while some others do not. The different frequency of SSWs occurrences may come from participants' personal manner of speaking, relationships between the participants, or the topic of the talk. Accordingly, 24 sets of conversation, 24 hours in total, are attended to and analyzed. No instances of SSWs are found in the rest of 7 sets. To provide the reader with the big picture of the data, Table 1.1 summarizes the background of each set.

Table 1.1
Description of Each Conversation

Data	Participants (Sex)	Place	Length (min)
BNT	Wada (F), Nakae (F), Tada (M), Suzuki (M)	Meeting room	28
BSK	Saki (F), Runa (F)	Participant's home	54
CHH	Nana (F), Akira (M)	Library study room	43
CHN	Abe (M), Yuki (F)	Classroom	54
DBT	Ken (M), Yuki (F), Nami (F)	Participant's home	58
DNS	Taku (M), Miki (F), Shun (M)	Library study room	65
FF	Shizu (F), Michi (F), Riko (F)	Classroom	35
GT	Ken (M), Tatsu (M), Yuki (F), Nami (F)	Participant's home	93
HKK	Tama (F), Nao (M), Mie (F)	Participant's office	47
KBN	Rui (M), Naoki (M), Azu (F)	Library study room	58
KK	Chie (F), Aya (F), Kana (F), Mako (F)	Participant's home	50
NYR	Kii (F), Rika (F), Sae (F)	Participant's home	100
OJS	Abe (M), Kato (M), Ogawa (M)	Meeting room	66
OKS	Yuko (F), Kimi (F)	Participant's office	47
OL	Aiko (F), Rina (F), Yumi (F)	Meeting room	46
OSK	Joo (M), Kei (F), Masa (M)	Library study room	43
OTR	Ako (F), Fumi (F), Rena (F), Yuka (F)	Library study room	55
SBT	Misa (F), Suzu (F)	Participant's home	24

SGK	Eri (F), Natsu (F), Kiko (F), Yuri (F)	Participant's home	144
SHR	Nao (M), Kaori (F), Yui (F)	Participant's office	40
SKR	Risa (F), Shota (M), Momo (F), Nao (M)	Library study room	46
SMM	Yasu (F), Moko (F), Ayaka (F)	Dining hall	96
WKR	Mio (F), Aki (F), Sanae (F)	Library study room	94
WSK	Hana (F), Aoi (F)	Participant's home	58
ZRG	Sae (F), Tomo (F), Rika (F)	Participant's office	77

From these 24 sets of conversation data, I have elicited the 313 target instances of SSWs in the following manner. I have reviewed conversation data, and first elicited 346 instances of SSWs. Then, I have excluded 33 instances that fall into the following four categories. First, I excluded phonomemies (*giongo* and *giseigo*), because they simply mimic the actual sound, such as *pippi↑ppi::* to mimic the time signal. Such onomatopoeias carry the simplest iconic pattern, which is not of interest in this study. Second, I excluded the instances of response cries (Goffman, 1978), which are "exclamatory interjections which are not full-fledged words" (Goffman, 1978:800). An example in Japanese is *eh* (Hayashi, 2009; Shimotani, 2008), which appear in my data in a quoted form (i.e., *eh tte kanji* 'like *eh*'). Goffman (1978) states that response cries "do not mark a flooding of emotion outward, but a flooding of relevance in" (p. 814). Thus, the concept of response cries does not exactly meet how I consider SSWs: sounds depicting manner of actions, states of objects, and people's emotional stances. Third, I excluded the instances of SSWs, produced by the speaker, who is out of the frame and is not visible in the frame, because participants' visible conduct carries great importance in this study. Finally, instances of SSW that are not fully audible, for example due to the background noise, are also excluded.

1.4.2 Transcription and English translation

The recorded conversation data was transcribed by hand by following Jefferson's (1985, 2004) transcription system. Transcripts are visible representations of interaction by means of alphabetical letters and symbols. Thus, the transcription conventions (p. viii) includes symbols that indicate, for example, the change of the voice pitch, length of pause, and the beginning of overlapping talk. In addition, snapshots of participants are inserted to the transcripts in order to make the bodily conduct accessible to the reader.

Japanese is a verb-final language. However, word orders are fairly flexible in spoken Japanese; thus, the verb may be followed by some elements, such as a noun or an adverb (Ono & Suzuki, 1992). In canonically structured sentences, each word is followed by a particle that marks the semantic relationship of the preceding noun and the predicate of the sentence.

For readers who are not familiar with the Japanese language, two lines are added to the transcribed interaction. The second line represents a word-by-word gloss. Abbreviations are listed in pp. viii-ix. The third line represents the English translation. Japanese is a verb-final language, while English word order is in principle subject-verb-object. The word order difference becomes an issue when it comes to translating Japanese conversation data into English, because while examining ongoing interaction, both the analysts and the readers take a close look at the way each turn unfolds. Thus, this dissertation presents English translations that are as natural as possible, while preserving how the original Japanese turns unfold. However, from time to time, coming up with an adequate translation may not be possible. Furthermore, the way the turn unfolds could be the key for the analysis. In such cases, two types of translations are presented. One of them is a translation that follows the way the original Japanese utterance unfolds. Such translation may not make sense to some of the readers who have little knowledge

of Japanese. Thus, the natural English translation is also provided.

The target items, SSWs, are frequently followed by the quotative particle *to* or *tte* and the verb, forming what I call a sound symbolic unit. Examples are *wasawasa: tte* , *suko::n tte* , and *byo::: tte*. In this study, those sound symbolic units are translated as 'like *wasawasa:*,' 'like *suko::n*,' and 'like *byo:::*' respectively, containing "like" that is often used in enactment to introduce the speaker's feeling (Fox & Robles, 2010).

1.4.3 Methodology

This study employs two analytical approaches to reveal two different aspects of SSWs. Part 2 (Chapters 2 and 3) addresses how depictions are embodied through the temporal deployment of various multimodal resources upon vocalization of SSWs. Thus, I analyze and describe how speakers construct turns that incorporate SSWs, while referencing various gesture studies and linguistic studies on SSWs. The analytical approach is primarily influenced by descriptive gesture studies, such as Kendon (1975, 1980, 1990, 2004) and McNeill (1992). The analyses, especially in Chapter 3, provide speculations about iconic meanings created through SSWs. It should clearly be stated that these speculations are made based on the previous findings on sound symbolism in Japanese.[7] Thus, this speculative approach does not confirm how participants of the talk treat the relationship between the phonetic designs of SSWs and their meaning.

Part 3 (Chapters 4 and 5) investigates what speakers can do through deployment of SSWs. The analytical approach taken in Part 3 is informed by Conversation Analysis (CA) (Sacks, Schegloff & Jefferson, 1974). CA has emerged from the study of social organization and aims to discover how human interactions are structured. Face-to-face interactions are considered orderly structured process through which each participant constructs practices, performs actions, and

[7] Studies on sound symbolism and iconicity will be outlined in Chapter 3.

engages in activities. The central concern of CA is how a turn is understood by recipients whose understanding is displayed in the next turn. The next turn is where co-participants display their understanding of the prior turn, thereby confirming action performed in the prior turn, as phrased as "a next turn proof procedure" (Sacks, Schegloff & Jefferson, 1974:728).

Some linguists have applied CA to linguistic analysis and established Interactional Linguistics that aims to disclose how people accomplish actions by adopting certain grammatical forms in certain sequential contexts (e.g., Couper-Kuhlen & Selting, 1996; Ford, 1993; Ford, Fox & Thompson, 2002; Hayashi, 2003a, 2004, 2005b, 2009; Mori, 1999; Ochs, Schegloff & Thompson, 1996; Selting & Couper-Kuhlen, 2001; Sorjonen, 2001; Tanaka, 1999, 2000, 2001). Through examinations of naturally occurring interaction, scholars have shed light on linguistic elements that serve as an interactional resource for participants. Fox, Thompson, Ford and Couper-Kuhlen (2012) summarize contribution of Interactional Linguistics to CA:

> Perhaps the most profound contribution of Interactional Linguistics to Conversation Analysis, and to the study of talk-in-interaction more generally, is the cross-linguistic orientation of the field of Linguistics, which embraces the differences in the linguistic resources and practices through which participants in different speech communities construct interaction. Many of the practitioners of IL [Interactional Linguistics] come from a typological or cross-linguistic background, and this awareness of linguistic differences has been brought into work on interaction in a growing number of languages (p. 732).

Japanese is one of such languages. Moreover, Japanese is known as one of the languages having a rich inventory of SSWs. As this study will show, participants of the talk creatively adopt existing SSWs, or even invent new ones. By revealing interactional uses of SSWs, this study contributes to a better understanding of actual uses of SSWs, which are now supposedly

observable in all world languages (Voeltz & Kilian-Hatz, 2001a).[8]

In face-to-face interaction, participants not only speak, but may move his/her hands, face, and/or the whole body. Bodily visual resources are as significant as one's language in performing actions (Ford, Thompson & Drake, 2012). The participants may even engage in some other activities, such as eating, drinking, or web browsing in some cases. What this means is that all the resources available to the participants in the space, such as artifacts, body, and gaze, operate in some way in interaction (e.g., C. Goodwin, 1984, 1986, 2000; Ford & Fox, 2010; Hayashi, 2005a; Mondada, 2006, 2012; Mori & Hayashi, 2006; Sidnel, 2006; Streeck, 1993, 1994; Streeck, C. Goodwin & LeBaron, 2011). Thus, this study takes not only the language but also bodily conduct into account, aiming at discovering how the participants coordinate such available resources and shape interaction.

1.5 Organization of the dissertation

The body of this dissertation consists of two parts. Part 2 lays the groundwork for analysis of SSWs in interaction. Chapter 2 highlights hand gesture occurrence during vocalization of SSWs. It first presents numerical data that show approximately eighty percent of creative and invented SSWs accompany hand gestures, while less than fifty percent of conventional SSWs accompany hand gestures. By analyzing instances of creative and invented SSWs, Chapter 2 shows how vocalization of SSWs accompany hand gestures, visually representing the content of the speech. Chapter 3 examines another multimodal aspect crucially involved in the deployment of SSWs: SSWs as vocal gestures. Particularly, it concentrates on the SSW *pon*, which can be vocalized in

[8] Childs (1994) states that "[a]lthough they [ideophones] constitute a robust word category in African languages, ideophones are relatively neglected and are rarely integrated into linguistic descriptions" (p. 178). In summarizing outcome of conference on ideophones held in 1999, Voeltz and Kilian-Hatz (2001a) report that "[i]deophones are found in many more languages than expected" (p.3).

various ways, for example, by repeating it and/or manipulating the pitch and/or rate. Previous studies have described SSWs as vocal gestures, but no study has clearly revealed what these vocal gestures refer to. My analysis will be the first to offer systematic analyses of conversation data that show how SSWs operate as vocal gestures in face-to-face interaction.

Part 3 builds on these preliminary analysis and reveals speakers' accomplishments through the deployment of SSWs in face-to-face interaction. Chapter 4 first demonstrates how speakers concretize objects that are absent from the interactional space by deploying SSWs. While vocalizing SSWs, the speakers juxtapose multiple multimodal semiotic resources, such as verbal elements, vocal quality, and hand movements. This chapter then shows that participants count on the *possibilities* when and the *limitations* of deploying SSWs. SSWs entail possibilities due to their capacity of depicting various referents. Conversely, participants may face limitations when the sequential environments have not been sufficiently built for the recipient to make sense of the depicted referents through SSWs. Finally, the chapter demonstrates that objects, concretized through deployment of SSWs, operate as provisional elements in word search activities (e.g., M. Goodwin, 1983). The subsequent chapter, Chapter 5, centers on SSWs serving as a resource to depict the speaker's emotional stance in a given situation. First, the chapter demonstrates that the speaker displays trouble with depicting certain emotional stances, and then demonstrates how the speaker depicts the emotional stances through deployment of SSWs. Second, adopting Prior's (2016) notion of "emotional scaling," the analyses demonstrate that SSWs serve as a readily available resource for the speaker to render emotional stances that rapidly change the intensity from one to the other. Finally, the last part of Chapter 5 demonstrates that the recipients are also capable of deploying SSWs when displaying their understanding of the prior speaker's depiction of emotional stances.

In the concluding chapter, Chapter 6, I discuss contributions of this study to our knowledge of SSWs, while summarizing the findings of this study. I go on further to discuss implications of this study connected with previous findings and how this dissertation could be expanded and further woven into SSW research across languages and disciplines.

Part 2

Preliminary to Analysis:

How Depictions through Sound Symbolic Words Are Embodied

Chapter 2

Sound Symbolic Words and Bodily Conduct

2.1 Introduction

This chapter primarily examines speakers' deployment of hand gestures, which visually depict the referent, while vocalizing SSWs. Gestures are "the movements of hands and arms that we see when people talk" (McNeill, 1992:1). In investigating the actual use of SSWs, some scholars in typological studies acknowledged that vocalization of SSWs is frequently accompanied by hand gestures (e.g., Childs, 1994; Kita, 1997; Kunene, 2001; Voeltz & Kilian-Hatz, 2001a). Furthermore, use of SSWs was considered a social phenomenon (e.g., Childs, 2001; Kunene, 2001). Despite these acknowledgements, the existing scholarship has not examined data taken from naturally occurring interaction, except for Dingemanse (2011, 2013[1], 2014). However, as Dingemanse (2011) pointed out, the accessibility to new technologies in recent years enabled us to examine participants' deployment of multimodal resources in interaction. Owing to the availability of technologies, this chapter demonstrates that gestures contribute to shaping the turn that incorporates SSWs depicting referents. It also demonstrates that hand gestures may accompany the speaker's movement of other body parts, such as the head and torso.

Since my dataset contains a large number of instances of SSWs, this chapter first presents a set of numeric data that show the frequencies of gesture occurrences across three types of SSWs. The purpose of presenting this data is to make visible that, on the surface, the degree of conventionality is related to the occurrences and non-occurrences of hand gestures during the

[1] Dingemanse (2013) examined gesture coupling in everyday conversation in Siwu. He concluded that "there *is* [emphasis in original] a particularly tight coupling between ideophones and iconic gestures in everyday speech" (p. 161), especially in narrative contexts. In my data, some participants deploy SSWs while engaging in telling activities, but other participants do so when engaging in other activities, such as word searchers (see Chapter 4).

vocalization of SSWs.[2] Dictionaries, published in 1996 and 2007, serve as the index of each SSWs' conventionality as mentioned in the Introduction. Table 2.1 below shows the number distributions of the three types of SSWs (i.e., conventional SSWs, creative use of conventional SSWs, and invented SSWs) and how frequently each type of SSWs co-occurs with gestures. It consists of six columns, and the four columns in the middle present the number of instances classified by gesture types. The gesture type classification is made based on whether the co-occurring gesture is a hand gesture or not. In my dataset, hand gestures are found to be the most outstanding type of gesture among other types. This is consistent with Kita's (1997) finding that SSWs very often co-occurred with hand gestures. Thus, the second and the third columns from the left in the table are combined when calculating the percentages.

Table 2.1
The Numbers Distributions of SSWs That Co-occur with and without Gestures

	With hand gestures	With hand, along with facial and/or body gestures	With facial and/or body gestures (no hand gestures)	With no gestures at all	Total
Conventional SSWs	28	5	4 (5.5%)	35 (48.6%)	72 (100%)
	33 (45.8%)				
Creative SSWs	99	37	23 (13.6%)	10 (5.9%)	169 (100%)
	136 (80.1%)				
Invented SSWs	29	26	10 (13.9%)	7 (9.7%)	72 (100%)
	55 (76.4%)				
Total	156	68	37	52	313

It is remarkable that less than fifty percent of conventional SSWs go along with hand gestures, while approximately eighty percent of creative and invented SSWs accompany hand gestures.

[2] It might be impossible to judge exactly how much a SSW is conventionalized at one point, because language uses have been changing over the years. Thus, in this study, I classify instances into three rough categories (i.e., conventional SSWs, creative use of conventional SSWs, and invented SSWs) by relying on dictionaries as the basis (see Chapter 1, Section 1.2, for the categorization).

This chapter examines gesture occurrences during the deployment of three different types of SSWs. It concludes the analysis with a segment that illustrates participants' conduct that confirms the numeric data. More precisely, in the final segment, the SSW *gatsun* is produced by two different speakers. The first occurrence accompanies a hand gesture that synchronizes with the beat of the SSW, while the second occurrence does not. The analysis shows that the first occurrence of the SSW establishes the recipient's understanding of the SSW. Thus, the second occurrence, without any hand gesture, reveals the speaker's understanding that the SSW is one that is shared among the participants. It is important to draw attention to the fact that, as the examination in this chapter shows, whenever such hand gestures are deployed, SSWs are always vocalized in a unique and distinct prosody. For instance, vocalizing involves prolonging the vowel and/or pronouncing part of it in a high pitch. The details of prosodic designs upon vocalizing SSWs will be examined in Chapter 3.

This chapter is organized as follows. Before presenting data, Section 2.2 outlines gesture studies from both cognitive and interactional perspectives. Section 2.3 then focuses on gesture occurrences upon deployment of creative SSWs and invented SSWs. It demonstrates that hand gestures visually represent what the speech affiliate depicts, and further demonstrates that SSWs are always vocalized in a unique prosodic design. By doing so, this chapter reveals that each component involved in deployment of SSWs contributes to iconic representation of the speech content. It then examines the segment in which the second occurrence of an SSW does not accompany a gesture or distinctive prosodic design, unlike the first occurrence. The analysis reveals that once the participants establish an understanding of the target SSW, the second occurrence of the SSW does not go along with any audio visual display. Before concluding this chapter, Section 2.4 touches on deployment of facial gestures, which are one of the significant

visual means for making sense of the turn in social interaction. Section 2.5 summarizes the finding of this chapter.

2.2 Gesture studies from cognitive and interactional perspectives

This study considers gestures as visible interactional resources; however, this is not the only view in previous studies. Let us briefly trace back how hand gestures have been studied. There are two prominent existing views on gestures, namely cognitive and interactional perspectives.

2.2.1 Cognitive perspectives

The central concern of gesture studies from a cognitive perspective is how gesture, speech, and thought relate to one another (e.g., McNeill, 1992). According to scholars focusing on human cognition, hand gestures are a manifestation of the speaker's mind by means of the shape of the hand and its movement. One of their core concerns is whether or not gesture is derived from the same cognitive system as speech. By taking a quantitative approach, scholars paid attention, for example, to timing of gesture and speech (e.g., Mayberry & Jaques, 2000; Nobe, 2000) and synchrony of gesture and speech (e.g., de Ruiter, 2000; Levy & Fowler, 2000). Within such a framework, scholars confirmed that gesture and concurrent speech convey coherent content to recipients (Alibali, Flevares & Goldin-Meadow, 1997; Goldin-Meadow & Sandhofer, 1999; Goldin-Meadow, Wein & Chang, 1992; McNeill, Cassell & McCullough, 1994). In the same vein, there is a claim that gesture and speech complement one another in the delivery of meaning (de Ruiter, 2000, 2003; Goldin-Meadow, McNeill & Singleton, 1996). According to this view, gesture and speech are confirmed to work as a unified system, which shapes meaning.

There is a line of research that describes gesture shapes and their movement, while investigating the relationship between cognition and speech (Kendon, 1975, 1980, 1990, 2004; McNeill, 1992; Scheflen, 1963, 1964). It is useful to consider how gestures are categorized depending on the shape, movement, and type of speech affiliate. One of the most influential studies of hand gestures is McNeill (1992). His research was theoretically grounded in what McNeill calls "Kendon's continuum" (p. 37), which differentiates types of hand movements (see Figure 2.1). The left end of the continuum is where both language and hand movements are integrated, while the right end is where only hand movements play a role. McNeill's work focuses on the left end (gesticulation).

Gesticulation→ language-like Gestures → Pantomimes → Emblems → Sign Languages

Figure 2.1[3] Kendon's continuum (based on Kendon, 1988, cited in McNeill, 1992:37)

Based on this continuum, McNeill (1992) introduced two classes of gestures: imagistic gestures, which consist of iconic and metaphoric gestures, and non-imagistic gestures, consisting of deictic gestures and beats. Among these classes, imagistic gestures convey some sort of image created in one's mind, while non-imagistic gestures do not. More precisely, an iconic gesture carries "a close formal relationship to the semantic content of speech" (p. 12), while a metaphoric gesture "presents an abstract idea" (p. 14).

Kita (1997) focused on iconic gestures; however, as Kendon (2004) pointed out, a hand gesture does not necessarily fall into just one of the four types (i.e., iconics, metaphorics, deictics, and beats), and may display mixed characteristics across types. Furthermore, as de Ruiter (2003) stated, metaphoric gestures are also considered to be iconic in that they depict what the speech

[3] Kendon's (1988) original work did not formulate a continuum, but discussed characteristics of these five types of hand movements one by one.

affiliate expresses. These remarks are consistent with Kita's treatment of gesture types in his later work published in 2002. That is, there is no mention of the term "iconic gesture" when referencing his own 1997 work. Instead, Kita uses the term *hyooshooteki jesuchaa* 'representative gesture' that covers both *byooshateki jesuchaa* 'depicting gesture' and *chokujiteki jesuchaa* 'deictic gesture.'

The absence of the term "iconic gesture" in Kita (2002) seems to be in line with what my data show. The dataset of this study includes SSWs that depict not only motions but also invisible concepts, such as the speaker's inner state. Accordingly, the types of gestures are not necessarily iconic but could be metaphoric and even pantomimic as well (cf., a creative use of the conventional SSW *pon*, presented in Chapter 3 on SSWs as vocal gestures). That is, regardless of the gesture type, the gesture shape and its movement visually represent the content of the speech affiliation. Thus, unless it is necessary, this study uses the term "gestures" generically to refer to any type of hand gestures.

Gestures are not merely random bodily movements but can systematically be analyzed in terms of their structures. Kendon's (1972, 1980) earlier works proposed that hand gestures unfold in three phases: preparation, stroke, and retraction. McNeill (1992) expanded Kendon's discovery by adding semantic aspects and establishing, in turn, that the stroke phase manifests the meaning of the co-occurring speech. Differently put, a gesture stroke is where the core meaning of the gesture is delivered. For this reason, the stroke phase obligatory occurs, while preparation and retraction may not occur.

2.2.2 Interactional perspectives

Scholars interested in human social interaction argue that hand gestures are visible resources for recipients to understand the ongoing turn that incorporates the hand gesture (e.g., Goodwin, 1979,

1981, 2000; Sidnell, 2006; Streeck, 1993). A Speaker's entire body can also be a part of hand movement. Even though people take for granted, a noteworthy fact is that participants' hands, moving in the interactional space, are visible to themselves in the talk. Streeck (1993) clarified that, since a gesture is a component of an utterance, it can be received and interpreted by the participants in the talk.

As Ford, Thompson and Drake (2012) stated, bodily visual resources are not secondary in performing actions; nor are languages secondary. For example, C. Goodwin (1979) showed that gaze is one of the significant visual resources for the speaker to design the sentence being produced. C. Goodwin demonstrated that even within a sentence the speaker modifies the turn design so that the utterance is oriented to the recipient of the talk, who is being gazed at. While examining participants' simultaneous deployment of multimodal semiotic resources, C. Goodwin (2000) revealed that human interaction is established by juxtaposed resources, such as language, gaze, hands, whole bodies and artifacts, which are all available to the participants within the entire interactional space. In the course of the interaction, participants design turns using these resources in a moment-by-moment manner, thereby making sense out of them.

Additionally, the participants of the talk may engage in some other activities, such as eating, drinking, or web browsing. This fact suggests that all the resources available to the participants in their shared space, such as artifacts, body, and gaze, operate through interaction. The following analyses, thus, take not only language and hand gestures into account, but also the body movements in order to capture the entire picture of interaction.

2.3 Creative uses of conventional SSWs and invented SSWs

2.3.1 With hand gestures and distinctive prosodic design

This section demonstrates that creative and invented SSWs are accompanied by hand gestures that visually represent speech content. In what follows, this section presents two sets of segments for each type of SSWs. The first two segments demonstrate how creative SSWs accompany hand gestures and are vocalized in a distinctive phonetic design.

The following segment is taken from a talk among four participants, but we focus on two participants Yuki and Nami. Prior to the segment, participants have talked about how clean and tidy Tatsu's (one of the participants) place was. On behalf of Tatsu, Yuki informed Nami that a stapler was always placed in Tatus's mailbox. Nami asked the reason, and in the following segment, Yuki explains the reason. During her explanation, Yuki deploys the SSW ↑*WAsawasa*. One of the definitions of the SSW *wasawasa* that seems to be close to what is described in the segment is this: *kusa ya ki no ha wo ookiku yurashite kaze ga fuku sama* "the manner in which the wind blows and grass and tree leaves sways" (Ono 2007:505).

(2.1) [GT.0321]

17 Yuki: *shimbun:: na, [posuto ni hairu yan.*
newspaper FP mailbox to put TAG
'newspapers:: go into the mailbox, right?'

18 Nami: [*ah:::::::::::hai*
oh I.see
'oh:::::::::: I see'

19 [*hai*
I.see
'I see'

Yuki

Figure 2.2

*

20 Yuki: → [*de shimbun te*,
and newspaper TP
'and the newspaper'

21 Nami: *h*[*ai*
yes
'yes'

Figure 2.3

Figure 2.4

Figure 2.5

Figure 2.6

Figure 2.7

* * * * *

22 Yuki: → [*koo*, [↑*WAsawasa: tte*=
like.this SSW QT

23 Nami: [*mo:*
so.like
'so like'

24 Yuki: → *=naru [yan.*
become TAG
'like, becomes like ↑*WAsawasa*, right?'

In line 17, Yuki starts explaining the reason why a stapler is placed in the mailbox, by mentioning a general fact: whenever newspapers are delivered, they are put into the mailbox. As soon as hearing the subject *shimbun* 'newspapers,' Nami claims her understanding of what Yuki has projected to produce, by producing *ah::::::::::hai hai* 'oh::::::::::I see I see' in lines 18 and 19. Yuki continues her explanation after Nami's claim of understanding. In line 20, while saying *de shimbun te* 'and newspapers,' Yuki initiates preparation of the forthcoming hand gesture: Yuki brings both hands in front of her body with the fingers all straight (Figure 2.2). In line 21, Nami continues displaying that she understands what Yuki is going to describe in relation to *shimbun* 'newspapers' by producing *hai* 'yes.' Yuki then continues by producing the proximate demonstrative *koo* 'like this,' which accompanies the spreading of her arms (Figure 2.3). Yuki then brings the two hands close (Figure 2.4). This movement synchronizes with the vocalization of the first sound *wa* in *wasawasa*. The vocalization of the rest of the SSW keeps accompanying hand gestures. When Yuki vocalizes the second mora *sa*, she extends both hands with her left hand more widely to the front (Figure 2.5). When vocalizing the subsequent *wa*, she then waves both hands but with the left palm more widely expanded (Figure 2.6). The final *sa* accompanies a wide movement in which she spreads both arms to the sides (Figure 2.7). That is, Yuki's hand movements visually depict how the newspapers spread around when reading. It should be noted that the size of Japanese newspapers are bigger than that of American ones. The page sizes of Japanese newspapers are generally 406 by 545 mm. The width is longer than American standards. For example, the dimensions of *New York Times* are 305 by 559 mm.

The vocalization of the SSW ↑*WAsawasa* is not flat; it includes some distinct characteristics. First, the whole SSW is produced in a higher pitch than the surrounding talk. Second, the first mora *WA* is produced loud. Additionally, the whole word is stressed. As described above, such vocalization of the SSW synchronizes the hand gestures that visually depict the speech affiliation. That is, the deployment of hand gestures is synchronized with the vocal production of the SSW. It should be noted that vocalization and prosodic design will be extensively examined in the next chapter, Chapter 3.

The next instance is also a creative use of conventional SSWs. The segment is from a conversation among three participants who used to take the same seminar class when they were college students. They have been talking about research each of them conducted when they were college students. Moko mentioned the word *horistikku* 'holistic,' which prompted Ayaka to discuss her study habits in the past, compared to the current situation (lines 12, 14 and 15). The SSW ↑*suko::n* appears in line 16, in which Moko agrees with Ayaka by describing how her knowledge was gone. The SSW *sukon* is defined in dictionaries. For example, Ono (2007) defines it as *mono nado ga kantan ni ochitari hamattari nuketari suru sama* 'the manner in which things easily drop, fit, or fall out' (p. 204).

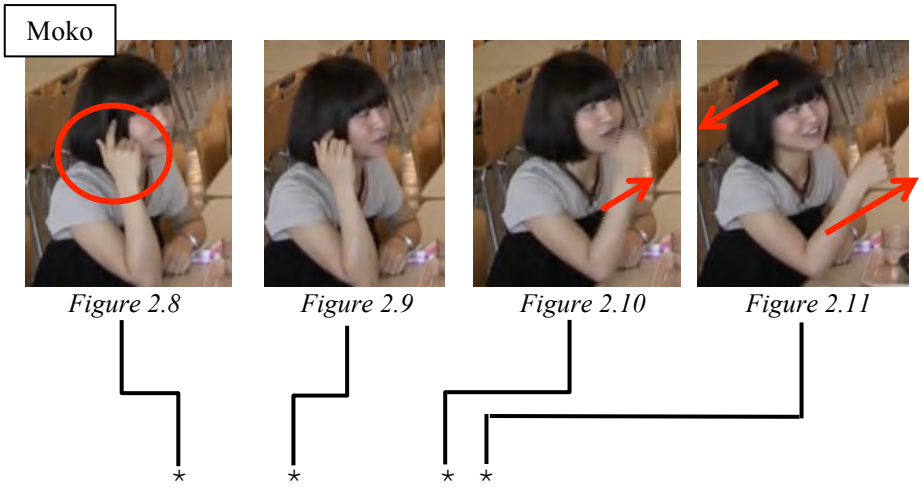
(2.2) [SMM.4647]

12 Ayaka: *nanka atashi ne, sono dai-, zemi de ne?*
like I FP like coll- seminar in FP
'like I, like at the coll-, in the seminar'

13 Moko: *un.*
'uh huh'

14 Ayaka: *benkyoo shiteta ninenkan*
study doing.PT for.two.years
'during the two years while I was studying there'

15 *dondake atama tsukatteta n daroo(h) tte o[mo(h)tte*
how.much brain using.PT N wonder QT think.and
'I think about how much I used to use my brain'



16 Moko: → *ya atashi mo nanka, su- ↑suko::n te nuketeru yo[huhhuh*
no I too like, su- SSW QT missing FP
'no, I do too, something is gone away like *su- suko:n*'

17 Ayaka: [ah so:
oh so
'oh you do?'

Responding to Ayaka's statement about her study habits in the past, Moko takes a turn that begins with *ya watashi mo* 'no, I do too,' indicating she agrees. Moko then produces *nanka* 'ike,' which indicates she is searching for an expression. She then produces *su-*, which is most likely the first mora of the subsequent production of the SSW ↑*suko::n*. The SSW is marked by the quotative particle *te*, which turns the SSW into an adverb that modifies the subsequent verb. That is, the SSW *suko::n* describes or even depicts how *nuketeru* 'something is gone away' (line 16).

The utterance does not specify the subject; however, the visual display indicates what is gone away and how. When Moko delays the turn by producing *nanka* 'like' (line 16), she brings her right index finger to the right temple (Figure 2.8). When the following cut-off *su-* is produced, she slightly moves the hand away from her head (Figure 2.9). While vocalizing the SSW

↑*suko::n*, she quickly moves the right finger to her left side across the face (Figures 2.10 and 2.11). Along with the hand gesture, her head also moves to the opposite direction of the hand (Figure 2.11). This series of body movements by Moko demonstrates that gestures are not limited to hands, but involves the whole body as well.

As stated earlier, when creative SSWs are produced, hand gestures frequently accompany them. Furthermore, when creative SSWs that accompany hand gestures are produced, the vocalization always carries certain characters. This phenomenon is seen in this segment as well. Moko vocalizes the SSW in a distinctive manner: First, the whole word is produced in a higher pitch. Second, the vowel *o* is prolonged as *suko::n*. This prosodic design depicts the speech content. Again, the vocal gestural aspect of SSWs will be examined in the next chapter.

The next two segments are from the collections of invented SSWs. Participants in the first segments are Yuko and Kimi, who are temporarily staying in the US where this conversation was recorded. In this segment, they are comparing varieties of snacks found in stores in Japan and the US. Thus, the demonstrative *kocchi* 'here' refers to the US. Kimi produces the SSW *byo::* in line 4 while presenting her observation of the amount of chips in stores in the US. The SSW *byo::* and variations are not found in dictionaries, thus this SSW is categorized as an invented SSW.

(2.3) [OKS.2121]

1 Kimi: *nanka ikko no shurui no ryoo ga, kocchi tte oo-*
like one.kind LK types LK amount SB here TP
'like, the amount of varieties of one kind of snack, here is *oo-*'

Figure 2.12 *Figure 2.13* *Figure 2.14*

2 *chippusu da to* (1.0) *nihon yor[i*
chips CP QT Japan than
'in case of chips (1.0) compared to Japan'

3 Yuko: [*u:n*=
'uh huh'

Figure 2.15

4 Kimi: → =[*byo::: tte aru ja nai desu ka.*
SSW QT exist CP NG CP Q
'there are like *byo:::*, isn't it'

5 Yuko: [*u::n u::n*
'uh huh uh huh'

In line 2, when Kimi (sitting in the right in the Figures) is about to finish mentioning the target of comparison (chips), she brings her hands up from the table to the height of her chest (Figure 2.12). She then keeps her hands apart across the width of her body, with her palms facing each other. At the same time, Kimi gazes at Yuko, whose gaze is directed at Kimi's hands (Figure 2.12). When the one-second silence completes, Yuko looks at Kimi and establishes a mutual gaze (Figure 2.13).

Kimi then produces the verbal component *nihon yori* ‘compared to Japan’ (line 20) and at the end of it, she starts preparing her hand movement (Figure 2.14) by bringing her hands closer to each other. This preparation overlaps with Yuko’s acknowledgement token together with a head nod produced in line 3. After receiving the recipient Yuko’s acknowledgement, Kimi produces the stroke of widening the space between her hands (Figure 2.15), in conjunction with the production of the SSW *byo::* (line 4). That is, the hand gesture is synchronized with the vocalization of the invented SSW. The production of *byo::* overlaps with Yuko’s continuation of the acknowledgement token and multiple head nods in line 5.

The next segment also contains an instance of an invented SSW that accompanies both sound emphasis and gestures, specifically, body movements that include hand gestures. In the previous section, we examined a segment in which the SSW accompanied not only a hand gesture but also head movement, suggesting that the whole body is also involved and works as a resource for the recipient to make sense of the prior turn that has incorporated the SSW.

The participants in the next segment are friends and are talking about pets. Sae has informed the recipients that she kept a hamster when she was a kid. However, according to Sae, it died *toire de* ‘in the toilet.’ Rika asked if it was *oborechattano* ‘drowned?.’ Sae denied it, then Rika, the other recipient Tomo, and Sae acted out one by one how the hamster would have tried to climb the toilet bowl. In what follows, Sae starts making a repair as *naka ja nakute* ‘not inside, and’ in line 26. The SSW ↓*eh:* is employed in line 28 in which Sae depicts how the hamster was found dead.

(2.4) [ZRG.2354]

26 Sae: *.hh .hh na(h)ka(h) ja nakute [:,*
inside at.TP NG.and
‘not ins(h)ide, and’

27 Rika: [.hh

Figure 2.16 *Figure 2.17*

28 Sae: → *nanka soto no kado de nanka, ↓eh: tte shindeta.*
like outside LK corner at like SSW QT dying
'like outside, at the corner, it was dead like ↓*eh:*'

As stated above, Sae deploys the word *naka* 'inside' in making a repair (line 26). She then specifies the exact spot using its antonym, which is *soto* 'outside,' in *soto no kado de* 'outside, at the corner' (line 28). Subsequently, by producing *nanka* 'like' (line 28), Sae indicates that the turn is continuing. Then she produces the SSW ↓*eh:* followed by the quotative particle *tte*, indicating that the preceding SSW was an adverb and projecting that the subsequent element will be a verb. As projected, the verb *shindeta* 'dead' follows. That is, the sound symbolic unit ↓*eh:* *tte* modifies how the hamster was found dead.

Upon the vocalization of the SSW↓*eh:*, Sae uses her upper body, including her arms. As Figure 2.16 shows, Sae does not make any particular movement up to the point in which she refers to the location *kado de* 'at the corner' (line 28). While producing the delaying item *nanka* 'like,' Sae starts moving her upper body. Sae arches her back and tilts her head to the front, and raises both hands to the shoulder level as if falling on her face. Sae momentarily pauses her body movement (Figure 2.17) during the subsequent vocalization of the SSW ↓*eh:* that is guttural and

produced in a lower pitch. Immediately after the temporal pause, her posture goes back and she sits still, closing the turn by producing the verb *shindeta* 'was dead.'

This section has demonstrated that when speakers vocalize creative and invented SSWs, they also produce hand gestures, which might further involve arm and body movements. Throughout the analyses, we have also observed that the vocalization of SSWs accompanies distinctive phonetic characteristics.

2.3.2 Without hand gestures nor distinctive prosodic design

The following segment contains two different versions of the same SSW *gatsun*, produced by two different speakers to describe food. The SSW *gatsun* is listed in the two dictionaries, published in 1996 and 2007, but the defined meaning is not related to food.[4] As will be demonstrated, the first vocalization of the SSW occurs together with prosodic and bodily visual resources, while the second verbalization of the SSW does not. This phenomenon suggests that, upon the vocalization of the second version of the SSW, the speaker treats the SSW as an already-shared item in terms of its meaning. In the segment, four colleagues in an office are talking about a new boss who prefers to eat something light for lunch. The target lines are elicited to show the differences between the two turns' multimodal designs.

In line 22 below, Suzuki informs the recipients of the reason why the boss does not go to the cafeteria in the office building: if they go and eat at the cafeteria, they have to eat like *gatsun*.

[4] Although the two dictionaries published in 1996 and 2007 do not have any food-related definition of *gatsun*, Super Daijirin Japanese Dictionary (2012) states that when *gatsun* describes an action of eating, it means the meal is rather heavy. This suggests that the food-related SSW *gatsun* is a relatively new expression.

(2.5) [BNT.0226]

22 Suzuki: *shokudoo de taberu to*
cafeteria at eat if
'if we go and eat at the cafeteria'

Figure 2.18

Figure 2.19

Figure 2.20

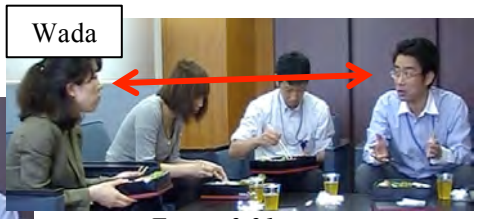


Figure 2.21

* * *

23 → *kekkoo gaTSUN to taberu kara:,*
pretty SSW QT eat thus
'we have to eat pretty *gatsun*'

At the beginning of line 23, while producing the adverb *kekkoo* 'pretty,' Suzuki brings both hands in front of his body and keeps the shape (Figure 2.18). When he verbalizes *ga* in *gaTSUN*, he moves both hands to the center of this body; both pinkies touch one another (Figure 2.19). While he verbalizes the second half of the SSW, *TSUN*, he brings both hands slightly upward, as if his hands were holding a bowl (Figure 2.20). One of the typical meals served in cafeterias in Japan is a big bowl of rice topped with meat, vegetables, and/or eggs, which is very filling. Suzuki's hand shape and movements iconically depict a big bowl in his hand. As Figure 2.21 indicates, while verbalizing the SSW, Suzuki's gaze is directed to the left, where Wada sits. It is not easy to see this in the snapshots, but Suzuki holds a pair of chopsticks in his left hand, while producing the gesture. It should be noted that he has been engaging in another activity, which is eating using the chopsticks. Prior to the elicited segment above, he was holding the chopsticks using both hands as Figure 2.22 shows.

Figure 2.22

That is, he changed the way he holds the chopsticks immediately before preparing the hand gesture that depicts the bowl.

After this turn, Suzuki continues informing the recipients that, due to the heavy meal served in the cafeteria, the boss goes to a convenience store and gets a light meal, such as *soba* noodles. As soon as the recipients hear this, they provides news receipt tokens such as *hee::* (Mori, 2006). Among the three recipients, Wada (sitting in the left), who received Suzuki's gaze in line 23, makes a remark about the impression of the boss in line 39 below. She articulates it through the same SSW *gatsun* in a different version.

(2.6) [BNT.0226-2]

Figure 2.23

39 Wada: → *nanka gatsun to tabesoo na fun'iki na noni*
like SSW QT eat.seem CP character CP though
'he seems to eat like *gatsun*, though'

In line 39, Wada verbalizes the SSW *gatsun* without any audible prosodic design. In addition, as Figure 2.23 shows, no hand gestures occur in the turn. One might consider that she cannot use her hands because she is grabbing the lunch box and chopsticks. However, first, we

observed that Suzuki shifted his way of holding the chopsticks and produced the hand gesture. There would have been a way for Wada to do the same. Second, the flat intonation upon the verbalization of the SSW reveals that the absence of the hand gesture is not necessarily due to the physical action of holding things but to her turn design. Importantly, while verbalizing the SSW, she gazes back to Suzuki, who first verbalized the SSW by utilizing multimodal resources of prosody and hand gestures. By bringing her gaze to him, Wada directs the remark toward Suzuki, who most likely shares the meaning of the SSW *gatsun* with her.

The second occurrence of the SSW *gatsun* in this segment provides us with a hint as to why vocalization of conventional SSWs may or may not accompany prosodic and visual designs. The non-occurrence of audio-visual practice reveals the participants' orientation toward the shared understanding of the SSW. Needless to say, conventional SSWs carry established meanings as defined in dictionaries; thus, they don't need to be accompanied by bodily conduct that visually pins down the meaning. In contrast, the first occurrence of the SSW *gatsun* might have been treated as the one that has not been fully conventionalized. Or, it might have been treated as the one that requires extra multimodal work to pin down the meaning; thus, the SSW might have been accompanied by hand gestures.

2.4 Notes on facial gestures

The analyses so far have focused on the speaker's bodily conduct, including hand gestures and body movements, upon the vocalization of SSWs. The vocalization of SSWs can also be involved with another multimodal resource, which is facial gestures. Facial gestures have been studied from an interactional perspective, and are recognized as visual resources that partially shape the interaction. For example, Peräkylä & Ruusuvuori (2006) and Ruusuvuori & Peräkylä

(2009) demonstrated that facial gestures ("facial expressions" in their terminology) provide clues about the stance taken by the recipient of the telling. In my dataset, facial gestures are deployed when the speaker acts out some sort of emotional stance (e.g., *gyo tte kanji* 'I feel like *gyo*'). Chapter 5 will examine cases in which the speaker makes use of a facial gesture upon the deployment of the SSW while depicting his/her emotional stance(s).

2.5 Summary of chapter 2

This chapter first presented a set of numeric data that showed the type of multimodal resources produced and its frequency upon the vocalization of SSWs. The data illustrated that creative SSWs and invented SSWs more frequently co-occurred with hand gestures than conventional SSWs. This chapter then demonstrated that, in actual face-to-face interaction, the vocalization of SSWs accompanies hand gestures that may involve movements of other body parts, such as arms and torso. Third, the last section analyzed the case in which two different speakers deployed the same SSW. It demonstrated that the second speaker displayed her understanding that the SSW was already a shared item and the meaning was established among the participants. This further confirmed the numeric data and explained why conventional SSWs less frequently co-occur with hand gestures. Finally, the overall analyses of this chapter revealed that vocalization of SSWs accompanying gestures is always delivered in a distinct prosodic design, such as in a high or low pitch and/or prolonging vowels. The vocal gestural aspect of the SSWs is the focus of the next chapter.

Chapter 3

Sound Symbolic Words as Vocal Gestures

3.1 Introduction

This chapter examines how sound symbolic words (SSWs) operate as vocal gestures in face-to-face interaction. Scholars have already described SSWs as "vocal gestures" (Voeltz & Kilian-Hatz, 2001a:3), although they have not clearly stated what the term "vocal gestures" refers to. Through examinations of moment-by-moment progression of turns, this chapter reveals that an SSW can be vocalized in various ways by changing the pitch and/or rate of speech. In the case of the SSW *pon,*[1] which is the object under study in this chapter, its phonetic shape can be *po::n*, >*ponponponpon*<, or <*POnPOnPOn*>, depending on what the speaker is depicting through his/her production of the SSW. Given this flexibility of vocalization, the speaker designs the phonetic shape of the word, thereby creatively gesticulating in face-to-face interaction.

Before proceeding, let us clarify that the words "gesture" and "gesticulate" are not exclusively related to body movement, but to the human voice as well. The most relevant study was conducted by Perlman (2010), who noted that words can be vocalized faster or slower than the surrounding talk, calling the phenomenon vocal gestures. No clear definition is provided, but Perlman seems to restrict vocal gestures to changes of the speech rate (e.g., fast or slow). In his experimental study, Perlman asked participants to describe scenes from a video segment that he had previously selected from the streaming website YouTube. He recorded the participants' re-telling of the scenes. When analyzing their accounts of the scenes, consisting of 142 individual descriptions, the speech rate was evaluated by measuring syllables per second. As a result, he

[1] Japanese SSWs are categorized into monosyllabic SSWs, or a CV base, and bisyllabic SSWs, or a CVCV base (Hamano 1998; Kakehi & Tamori 1993; Tamori, 1993).

found that "fast adverbials were spoken significantly faster than slow adverbials, with means of 12.86 phones/sec and 10.54 phones/sec" (p. 10). One of the utterances the participants produced is: "There's another girl bouncing on a ball really ssssloowly…" (p. 12). According to Perlman, the final adverb "slowly" in this example contains two significant elements. These are the word initial *sss* and the vowel lengthening *oo*, which convey speech content through its slowness. Through these elements, the speaker sound-iconically depicts the speech content. Speech rate is confirmed to be iconic, representing the speech content.

While analyzing the actual use of ideophones in interaction, Dingemanse (2011) highlighted the same phenomenon in his arguments of depicting power of ideophones. One of his cases demonstrated that the speaker's actual production of the SSW captures the timing of the pounding event the participant depicts in talk: *ǹdo lowètè* (0.2) ↑*kpɔ*↑ (0.9) ↑*kpɔ*↑ (0.9) 'I'm pounding, (0.2) ↑*kpɔ* ↑ (0.9) ↑*kpɔ* ↑ (0.9)' (p. 177). Although Dingemanse did not use the phrase vocal gestures, he argued that repetition of an ideophone is not merely repeating an event but "an accurately timed recreation of the pounding event" (p. 175). These two studies have influenced my views on the vocal gestural aspects of SSWs in interaction. Thus, in this study I define "vocal gesture" as follows:

> A form of vocal representation in which the speaker may creatively design the phonetic shape of the word and vocalizes the word as designed, thereby acting out the referent.

This definition is formulated to describe the core feature of vocal gesture, while acknowledging the differences between conventional SSWs and invented SSWs when SSWs operate as vocal gestures. Vocal gestures are a form of vocal representation. Like any other vocalized words in interaction, the vocalized words become audible resources for recipients to make sense of the utterance. What is significant about vocal gestures is that vocalized words convey concrete

speech content by means of distinctive phonetic designs. SSWs constitute a class of lexical items that shows a non-arbitrary relationship between speech sound and meaning. Thus, the vocalized word directly links to its own descriptive meaning. With the help of the descriptive efficiency that characterizes SSWs, the speaker creatively designs the phonetic shape of the word. He/she can freely manipulate aspects of vocalization, such as rate and pitch, through which the speaker acts out the referent in the current site of the interaction. Significantly, not only vocalization of SSWs is iconic. Production of SSWs are often accompanied by hand gestures (Kita, 1993, 1997), whose shape and movements, in many cases, iconically designate speech content as demonstrated in Chapter 2.

The purpose of this chapter is twofold. First, this chapter aims to reveal how the characteristics of vocal gestures are embodied. The analyses in this chapter demonstrate that speakers make use of the flexibility of vocal production upon the vocalization of SSWs by manipulating their pitch and/or rate. By vocalizing SSWs, speakers act out objects that are not available in the current site of the interaction. The analyses also demonstrate that vocal production is accompanied by visible hand gestures. By coordinating various semiotic resources, speakers audibly and visually design the turn that incorporates the SSW, where gestural characteristics come into being. Second, this chapter aims to show how deployment of SSWs serve multifaceted resources for the speaker to depict the referent in face-to-face interaction. By analyzing the design of each turn, this chapter reveals that speakers are capable of depicting various types of referents, from physical motions to invisible concepts.

In order to conduct a systematic analysis, this chapter scrutinizes conversation segments, all of which consist of one or more sounds of *pon*. This chapter concentrates on the sound *pon* due to the availability of data segments. Of more than 340 instances of SSWs in my database,

pon is the most versatile sound component. In total, 8 instances are elicited and all of them are examined in this chapter. Among these *pon*-containing SSWs, two of them are conventional uses of SSWs, while all the rest constitute creative departures from the conventional SSW. The sound *pon* consists of two moras (i.e., *po* and *n*) that can be phonetically designed in various ways, for example, *pon pon*, *po::n*, and *> pon pon pon< ↑po::n*.[2] That is, the vocalization of the word *pon* can be flexibly manipulated in terms of rate and pitch, as well as in the number of times the sound is repeated. Moreover, the vocalization of SSWs co-occurs with hand gestures, which visually and iconically display speech content (e.g., McNeill, 1992). The analyses demonstrate that, by means of the flexibility of vocalization intertwined with concurrent hand gestures, speakers depict various types of referents, such as visible motions, invisible concepts, and intangible objects in the current site of the interaction.

Furthermore, although the focus of this chapter is SSWs as vocal gestures, the close examinations of each segment turns out to show interesting and significant features of gesture occurrences, upon the vocalization of SSWs. That is, when the referent of the SSW description is clear to the recipient at the moment of vocalization, the gesture synchronizes with the beat of the SSW. In contrast, when the referent of the SSW description is not clear to the recipient at the moment of vocalization, the co-occurring gestures do not necessarily rhythmically synchronize with the sound beat of the SSW. Analyzing data taken from a re-telling task in Japanese, Kita (1993) found that "when mimetics are accompanied by a stroke, the mimetic and stroke tend to be accurately synchronized" (p. 91). In his quantitative analysis, 94% of mimetics were synchronized with a gesture stroke. My qualitative, moment-by-moment, analysis of naturally-occurring conversation data will shed light on the hidden side of gesture use, namely, non-synchronization.

[2] These examples are taken from segments that will be examined later in this chapter.

The remainder of this chapter is organized as follows. Before analyzing actual data, Section 3.2 outlines previous studies on sound symbolism and iconicity. Section 3.3 then explores how scholars in linguistics have established the meaning of the SSW *pon*. This is followed by data analysis, demonstrating how the SSW *pon* transforms itself in terms of the phonetic shape and meaning. Section 3.4 demonstrates that, while vocalizing the SSW, speakers manipulate their rate and pitch, thereby depicting motion. Section 3.5 then demonstrates that even invisible concepts can be acted out via phonetically designed SSWs with the aid of visual display through hand gestures. Section 3.6 focuses on SSWs that depict arrangements of motionless objects that are not available in the interactional site. The vocal representation of SSWs act out intangible objects arranged in or across a certain space with the help of hand gestures that visually display how the objects are arranged. Finally, Section 3.7 concludes this chapter by summarizing the findings of analyses.

3.2 Sound symbolism and iconicity

Previously, the existence of iconic relationship between the form and meaning has been taken for granted in SSWs studies. The concept of iconicity is concerned with the resemblance between the language form and its meaning (e.g., Hamano, 1998). Iconicity goes against the linguistic concept of sign that consists of a signifier and a signified, proposed by Ferdinand de Saussure, one of the important figures in structuralism. For example, in Japanese, the series of sound /*i nu*/ refers to what the English word dogs refer to, but there is no iconic relationship between the word form and its referent/meaning. In studying SSWs, scholars have sought to find regularities that associate the sound symbolic feature with the word meaning. For example, there seems to be a consensus that /Q/ at the end of a word symbolizes a sense of quickness and a sense of speed

(Tamori, 1993), while a word-internal /Q/ conveys emphasis. It is also widely known that a word-final /N/ conveys a sense of resonance (e.g., *batan*). Word-internal long vowels often represent something physically long in the real world (e.g., *baan*). Importantly, however, Dingemanse (2001), through examination of ideophones in Siwu, argued that not all the ideophones are clearly iconic.[3] He further demonstrated that the primary role of ideophones is not iconically designating the speech content, but depicting the referent, arguing that "the power of depiction in speech can be experienced in full force" (p. 176). This study deals with the same phenomenon under the category of SSWs as vocal gestures.

In this study, the multimodal analysis of naturally occurring Japanese data will show that the premise held by previous studies about existence of iconicity goes hand-in-hand with Dingemanse's claim introduced above. The overall picture that emerges from the analyses of this study is that all the multimodal resources, which co-occur with or precede SSWs, are interlaced in shaping the meaning and/or recipients' understanding of the SSWs in the ongoing face-to-face interaction. More precisely, as Kita (1993, 1997) clarified, hand gestures frequently co-occur with SSWs and iconically depict the speech content (Chapter 2). Furthermore, upon the deployment of SSWs, the speaker's voice also iconically gesticulates. The analyses will demonstrate that the speaker depicts the referent by juxtaposing various multimodal resources upon deployment of SSWs. Even when the form of SSWs is not transparently iconic, at least one of the co-occurring semiotic resources appears to be iconic.

Within the studies of Japanese SSWs, Tamori and Schourup (1999) analyzed characteristics of vocalizing *rinjigo* 'temporary SSWs,' which are equivalent to what this study calls invented SSWs. Tamori and Schourup stated that when speakers vocalize temporary words,

[3] Dingemanse inspects various kinds of iconicity (i.e., imagic iconicity, diagrammatic iconicity that includes gestalt iconicity and relative iconicity) to ensure his arguments.

they change voice quality. The voice can be breathy voice or creaky voice, or the length of the vocalization may be altered (p. 193). Another note made by Tamori and Schourup, which is rather intriguing, involves the following example sentences from English[4]: "The gold nugget went flash" and "The fire-bomb went flash." They argued that, the first sentence "The gold nugget went flash" is not grammatically appropriate. But if the second sentence is pronounced as "The fire-bomb went FLASSH!!!" in which the word flash is vocalized energetically, for example by breathy voice or by pronouncing strongly, the sentence serves as an appropriate expression. Tamori and Schourup further argued that vocalizing the word "flash" in a particular way turns the word into a sound mimicking word, or an onomatopoeia (p. 110). This statement shares the common core of vocal features with Perlman's (2010) study: Vocalization has the power to depict the referent as proximate to the form/movement/texture of the referent as possible. That is, human voice serves as an interactional resource for the participants when sound-symbolically designing the turn.

3.3 Meanings of SSW *pon* in previous studies

Prior to presenting the data analyses, it is useful to explore how previous studies have identified meanings of the SSW *pon*. This section first surveys its dictionary meanings, which are followed by a clarification of the phonetic meanings of /p/, /o/, and /N/. This is in turn followed by a summary of morphological and phonetic modifications of words as well as an explanation of the meanings generated by these modifications.

To begin with, let us investigate the dictionary meanings of *pon*. For Kakehi, Tamori, and Schourup (1996b), *pon* imitates "[a] small popping sound; a light clapping or patting sound" (p. 994). Ono (2007) provides two definitions of *pon* as sound mimetics (i.e., onomatopoeias,

[4] Tamori and Schourup made this argument only in discussing English data.

giongo, or phonomimes). The first definition[5] is "the sound that describes the manner of lightly throwing something, something popping out, or something bouncing" (p. 464). The second definition[6] is "the manner of responding instantaneously" (ibid). Ono (2007) defines *pon* as manner mimetics (i.e., *gitaigo* or phenomimes) as: "the high pitched sound when something bursts or pops, or its manner."[7] The second definition[8] is "the high pitched sound produced when things are hit, or its manner" (p. 464). *Pon* as manner mimetics is also defined in a different dictionary, as the sound that describes "[t]he manner of throwing, being thrown, landing etc., effortlessly or casually" (Kakehi, Tamori & Schourup, 1996b:995). Hamano (1998), who conducted an extensive study on Japanese sound symbolism, provides contexts where *pon* can be used. The dictionary definitions introduced above overlap with the contexts[9] provided by Hamano: "tapping the shoulder or a hand drum; the manner in which a heated chestnut cracks; clapping hands lightly without tensely stretching the fingers; throwing a ball casually" (p.78). These attempts to define *pon* do not necessarily remain elusive, but the list seems to go on.

Having explored these dictionary definitions, let us scrutinize these in depth by examining the sound symbolic meaning each component of *pon* carries, namely, the word initial labial stop /p/, the vowel /o/, and the word-final nasal /N/. According to Hamano (1998), the initial consonant /p/ carries the meaning of "an abrupt and explosive movement or event such as hitting and explosion, or the precondition for such in the form of tension" (p. 86). The vowel /o/ conveys the meaning "where the affected area or object is small or where the event affects only part of the object" (p. 77). The word-final /N/ "indicates that the direction of the motion or the

[5] The original Japanese wording is: *karuku nagedashitari tobidashitari hazundari suru sama*.
[6] The original Japanese wording is: *shunkanteki ni oojiru sama*.
[7] The original Japanese wording is: *mono ga haretsu shitari ikioiyoku nuketari suru toki no takai oto. Mata sono sama.*
[8] The original Japanese wording is: *mono o tataku karui takai oto. Mata sono sama.*
[9] Although Hamano uses the term "context," the list she provides has not only contexts but more concrete and specific scenes in which these SSWs can be used.

quality of the sound changes toward the end. That is, it indicates that there is a reaction to or reverberation after the initial movement. It may indicate that the object is flexible or elastic and is capable of such reactions or reverberations" (p. 67). In sum, as much as the dictionary definitions of *pon* vary, all entail these phonetic meanings.

Given the meanings and associations of the sound *pon*, I now outline what kind of effects are sound-iconically conveyed when the shape of *pon* is modified by means of repetition, duration change, and rate change respectively. The notions of duration and rate are similar in that both are linked with the span of vocalization. In this study, when part of the word is lengthened or shortened compared to the original dictionary pronunciation, such a modification is considered an instance of change in duration (e.g., *po::n*). Conversely, when the whole word is produced at a faster or slower speed than the surrounding talk, the modification is considered a case of rate change (e.g., >*ponponpon*<).

First, the SSW *pon* can be repeated as *pon pon, pon pon pon, pon pon pon pon*, and so on. The number of times an SSW is repeated "closely parallels the actual number of acoustic signals or actions" (Hamano, 1998:65). An example is cited below. It should be noted that Hamano's examples are associated with *pin*, not with *pon*.

(3.1) a. Teeburukurosu-o piN-to hiQpaQ-te hosita.
table cloth-Acc pull-and dried
'I carefully pulled the table cloth into shape and hung it on a line.'

b. Teeburukurosu-o piN piN-to hiQpaQ-te hosita.
table cloth-Acc pull-and dried
'I carefully pulled the table cloth into shape a couple of times and hung it on a line.'

(Hamano, 1998:65)

In (3.1a), the action of pulling is done once, as the occurrence of *piN* shows. In contrast, in (3.1b), the action of pulling is done twice, as the sequence *piN piN* demonstrates.

Secondly, in the Japanese phonological system, all words consist of at least one vowel, of which the vowel can be lengthened. When a vowel is lengthened, this lengthening expresses that "the action takes longer spatially or temporally and is more strenuously carried out" (Hamano, 1998:72). The following examples illustrate this point.

(3.2) a. PiN-to hiQpaQte.
stretch
'Stretch it tight.'

b. PiiN-to hiQpaQte.
stretch
'Stretch it very tight.'

(Hamano, 1998:72)

In (3.2a), which contains the original version *PiN*, the action of stretching is done tightly, as indicated in the English translation. In contrast, in (3.2b), which contains the prolonged version *PiiN*, the action of stretching is done very tightly.

Taking into account these previous findings regarding sound symbolism and iconicity, the rest of this chapter examines how SSWs operate as vocal gestures in naturally-occurring conversations in Japanese.

3.4 Motions

This section examines three cases of *pon*-containing SSWs, all of which depict motions. The first two cases (i.e., >*ponponpon*<↑*po::n* and <*ponponpon*>) concern physical motions; the last case (i.e., *pon*) contains an SSW that depicts the motion of an object. In the first two cases, speakers phonetically design the vocalization of SSWs by repeating the sound *pon* and manipulating the pitch and the rate in vocally acting out certain motions. This is where the characteristics of vocal gestures appear in interaction. In contrast, the third case of *pon* (i.e., depicting an object's

motion), does not carry any distinctive phonetic features, due to the nature of the motion the speaker depicts through the SSW. However, the examination shows that all the vocal representations in the three cases are coordinated with the speaker's body movements, thereby depicting the motions of each respective referent.[10] Also, gestures are rhythmically synchronized with the beat of SSWs. Each occurrence of the SSW embodies its meaning in the individual interactional context.

Let us examine the first segment. Prior to the segment below, Saki has informed Runa that she went to see a basketball game and saw a "buzzer beater." Saki displays she was not sure if it was called a buzzer beater, so she asks Runa: *buzaa biitaa da kke* 'was it called buzzer beater?' Here, the final particle *kke* indicates she was unable to remember the name and sought help Runa to recall it (Hayashi, 2012). While seeking Runa's help, Saki makes the gesture of dribbling with her right arm. Saki then comes to realize it is *buzaa biito*. When Runa confirms it by producing *un* 'uh huh,' both Saki and Runa come to understand the topic at hand. In the following segment, Saki vocally acts out the player's motion, as well as the motion of the ball, by means of vocalizing the SSW, *>ponponpon<↑po::n*, in line 6.

(3.3) [BSK1.4503]

4 Saki: *ko yatte*
like.this do.and
'doing something like this'

5 *isshoni hashitteru dake na n da kedo:*
together running only CP N CP but
'players are just running along, bu:t'

[10] Occurrences of SSWs, especially those that depict motions, can be considered reenactments of past events. Sidnell (2006) explains that "reenactments are like direct reported speech and demonstrations in that they depict or show rather than describe" (p. 381). The key word here is "depict," because the deployment of SSWs also accomplishes depiction through interaction. However, the focus of this chapter is how SSWs operate as vocal gestures. Thus, I am not going to consider how the deployment of SSWs accomplishes a reenactment. A further systematic analysis will confirm the relationship between the deployment of SSWs and reenactments.

6 → *hontoni >ponponpon<↑po::n te nagetara haitte:*
really SSW QT throw.when enter.and
'when a player throws a ball like >*ponponpon*<↑*po::n*, the ball went in the net'

In lines 4 and 5, Saki explains the players' motion: *ko yatte isshoni hashitteru dake na n da kedo:* 'doing something like this, players are just running along, bu:t.' By marking the end of the utterance with *kedo* 'but,' Saki indicates that her explanation of the players' movement continues. Saki then produces *hontoni* 'really' (line 6), and vocalizes the SSW as >*ponponpon*<↑*po::n* in line 6. This variation of *pon* consists of four occurrences of *pon*, the first three of which are produced consecutively in a fast rate. The number of times Saki repeats the sound *pon* indicates that the motion associated with *pon* occurred three times. Furthermore, the accelerated vocalization >*ponponpon*< indicates that the motion was completed in a short time. What immediately follows is ↑*po::n*, in which the vowel /o/ is lengthened. This phonetic design sound-iconically indicates that completing one motion takes time. Additionally, the SSW is vocalized in a higher pitch. The vocalization in a high-pitched sound might sound-iconically indicate that the motion acted out by *po::n* physically occurred in a high position. That is, Saki designs her turn to display the player's motion of dribbling and shooting a basket, which matches dictionary definitions. The meaning of the first three occurrences of *pon* corresponds to the dictionary definition of "something bouncing" (Ono, 2007:464), and the last occurrence of lengthened version *po::n* "throwing something" (ibid). Importantly, the immediate prosodic change differentiates the referent of *po::n* from that of >*ponponpon*<. The speaker makes use of the flexibility of the vocalization of *pon* through its repetition and modification of the rate and the pitch, thereby vocally acting out the player's motion.

Grammatically speaking, this SSW >*ponponpon*<↑*po::n* is marked by the quotative particle *te*, followed by the verb *nageru* 'to throw.' That is, the SSW functions as an adverb that modifies the verb that follows. Given that the following verb is *nageru* 'to throw,' the final portion of the SSW ↑*po::n* clearly displays the manner in which something is thrown. The first portion of >*ponponpon*< does not necessarily relate to the manner of throwing something, because, as mentioned, it is detached from the final portion by the prosodic change.

A close examination of Saki's turn reveals that the vocalization of the SSW is synchronized with her deployment of hand gestures and gaze movement. By coordinating all available resources, Saki makes the meaning of the SSW audible and visible, thereby shaping the meaning of the SSW in face-to-face interaction.

(3.4) [BSK1.4503]

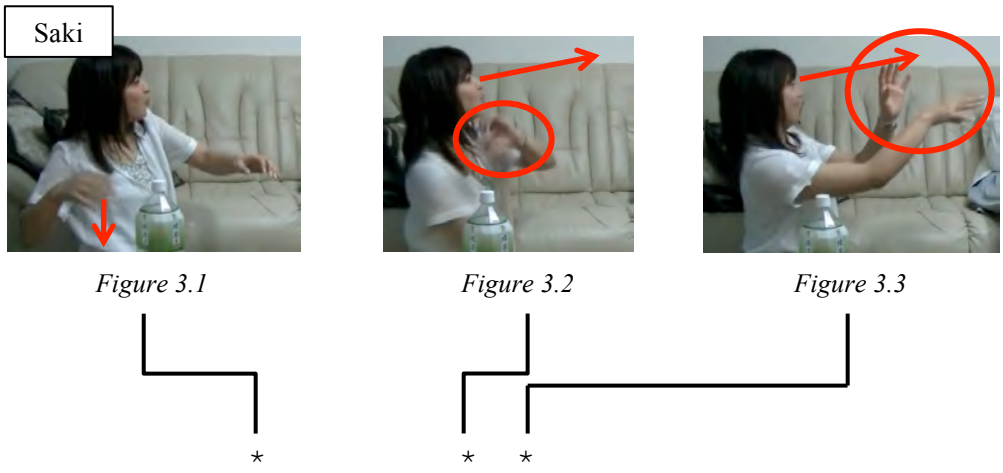


Figure 3.1 *Figure 3.2* *Figure 3.3*

6 Saki: → *hontoni >ponponpon<↑po::n te nagetara haitte:*
really SSW QT throw.when enter.and
'when a player throws a ball like >*ponponpon*<↑*po::n*, the ball went in the net'

While producing *hontoni* 'really' (line 6), Saki starts to prepare for a gesture with her right arm. Immediately before starting to vocalize the SSW, she brings her right arm next to her body, with

the palm facing down. Then, she imitates the quick movement of dribbling the ball with her right arm (Figure 3.1). This movement is rhythmically coordinated with the vocalization of the first three occurrences of >*ponponpon*<. In other words, each occurrence of *pon* is synchronized with the body gesture that mimics the player's dribbling the ball once. Immediately after the third occurrence of *pon*, Saki switches her gaze and stares at a fixed point ahead of her. She then moves both hands to the front of her right shoulder (Figure 3.2). At this moment, the palms are facing one another as if she were holding a ball. Saki then vocalizes the fourth *pon* as ↑*po::n* in a high pitch. This vocalization is still coordinated with her body movements. While vocalizing ↑*po::n*, Saki enacts the movement of throwing the ball forward with her arms (Figure 3.3). Her gaze is supposed to aim at the spot where she is throwing the imaginary basketball. By making use of the flexibility of the vocalization, along with the deployment of hand gestures and gaze direction, Saki depicts a player's motion as it occurred somewhere else in the past.

The next segment takes place approximately seven minutes after the previous segment. The participants are still talking about basketball games. In this segment, the other participant, Runa, produces the *pon*-containing SSW <*POnPOnPOn*> in line 11, in which Runa depicts how a player moves from one side of the basketball court to the other.

(3.5) [BSK2.0553]

6	Runa:	*s: suggoi nanka attoyuumani sa* very like quickly FP 'really quickly'
7		*hashi kara hashi made [sa* end from end to FP 'from one place to another'
8	Saki:	[*un* 'uh huh'

9 Runa: *modot chau kara sa:*
return done thus FP
'players return, so'

10 Saki: ((nodding while eating))

11 Runa: → *shikamo <POnPOnPOn> mitaina sa:*
also SSW like FP
'also, like <*POnPOnPOn*>'

In lines 6, 7 and 9, Runa explains how the players return to their original positions as *suggoi nanka attoyuumani sa hashi kara hashi made sa modoc chau kara sa:* 'really quickly form one place to another, players return, so.' The recipient Saki, who is eating snacks, provides Runa with a small nodding for acknowledgement. Runa then continues her turn with *shikamo* 'also' (line 11), which indicates that she is going to provide Saki with additional commentary on how players returned to their original positions. At this moment, the SSW <*POnPOnPOn*> occurs.

The close examination of how the speaker vocalizes the SSW reveals the way the SSW operates as a vocal gesture. First, three occurrences of *POn* are produced in a row without breaks. Second, as the transcription symbol shows, the SSW as a whole is vocalized more slowly than the surrounding talk. That is, the voice iconically displays the motion taking place slowly. Furthermore, each *PO* in *POn* is strongly emphasized, thereby sound-iconically indicating the heaviness or massiveness of the movement. In addition, the SSW is followed by the word *mitaina* 'like,' which grammatically indicates that the preceding SSW has related to a way of performing an action. By virtue of this particular position in the turn, the SSW appears to act out the manner in which the players return to their positions on the basketball court with three steps, possibly with the stride as sound-iconically indicated by the slow and stressed vocalization of <*POnPOnPOn*>. In other words, the SSW operates as a vocal gesture acting out the way in

which the players walk on the basketball court from one side to the other. It should be noted that this use of *POn* is not related to the conventional meanings surveyed in Section 3.3. Here, the SSW *POn* is deployed as a resource for depicting the player's movement. Saki's and Runa's distinctive manipulation of the pitch in vocalizing the SSWs (from Segments 3.3 and 3.5 respectively) are shown in the following Figures.[11, 12]

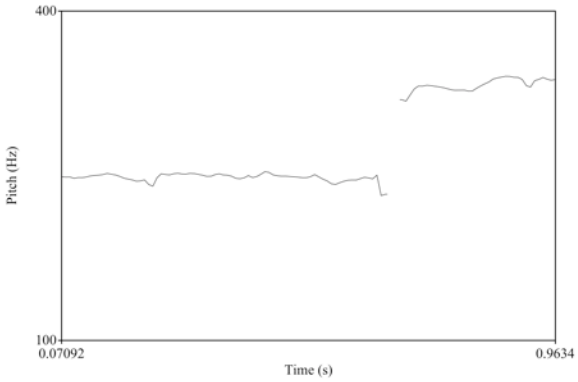


Figure 3.4. *Segment 3.3 line 6, Saki:* >ponponpon<↑po::n

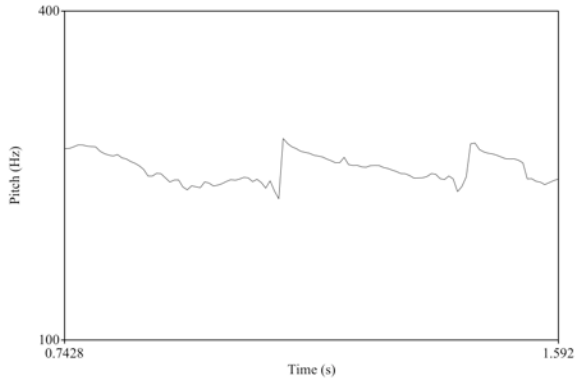


Figure 3.5. *Segment 3.5 line 11, Runa*: <POnPOnPOn>

[11] The figure is formulated through the acoustic analysis software *Praat*. For further information, see http://www.fon.hum.uva.nl/praat/.

[12] Unfortunately, due to the overlapping talk or background noise, the acoustic analysis of pitch contour of other *pon*-containing words cannot be available for the rest of the analyses in this chapter.

Not only does vocalization contribute to embodiment of the SSW meaning, but the accompanying body movements also contribute to it.

(3.6) [BSK2.0553]

Figure 3.6 *Figure 3.7* *Figure 3.8*

* * *

11 Runa: → *shikamo* *<POnPOnPOn>* *mitaina* *sa:*
also SSW like FP
'also, like <*POnPOnPOn*>'

In line 11, while Runa produces <*POnPOnPOn*>, the sound beat is rhythmically coordinated with her arm movements. She swings her arms widely, and each swing is in concert with the vocalization of one *POn* sound (Figures 3.6 to 3.8). What is displayed by this wide body movement coincides with the sound representation in which *PO* in *POn* is stressed.

These two segments contained SSWs that depict human physical motions. The next segment is also for examining a case of motion, but of an object. Again, the next segment shows that the vocalization of the SSW is rhythmically coordinated with body movements. However, unlike the previous two cases, no audibly unique phonetic design is observed. This non-occurrence of prosodic change is probably because the motion depicted through the SSW involves a small motion that is accomplished in a short period of time. In addition, what *pon* depicts in this segment corresponds to the dictionary meaning of "popping sound" (Ono,

2007:464). The conventional meaning that *pon* carries might be related to the non-occurrence of prosodic change: The speaker does not need to pin down the meaning by means of iconic vocalization. In other word, the speaker might have designed the phonetic shape of the SSW as it was in order to display that the SSW did not emerge creatively but conventionally.

In the following segment, Kei is describing how a capsule-toy comes out of the vending machine. A little while ago, Kei started explaining the procedure of using the capsule-toy machine. The first part of the explanation was not printed in the transcript. The following segment takes place in the midst of her explanation, in which the way the capsule-toy comes out of the vending machine is depicted through the SSW *pon* in line 17.

(3.7) [OSK.4130]

17 Kei: → *atattara pon te,*
win.if SSW QT
'if you win, like *pon*'

((1 line of the recipient's response is omitted.))

19 Kei: *nanka gachagacha detekun nen*
like capsule-toy come.out FP
'like, the capsule-toy comes out'

In line 17, the initial element is *atattara* 'if you win,' which indicates that the subsequent element describes the result of winning the game. Right away, Kei produces the SSW *pon*, followed by the quotative particle *te*. Since Kei has been engaged in explaining the procedure of how to use the machine, the SSW is heard as the consequence of winning the game—getting a capsule-toy. As she indicated during the production of her turn, Kei continues her turn as *nanka gachagacha detekun nen* 'like, the capsule-toy comes out' (line 19).

In the vocalization of the SSW *pon* in this segment, there is neither sound lengthening nor pitch change. Moreover, there is no repetition of *pon*. In this way, the vocalization of *pon* operates as a vocal gesture. Here, the motion depicted via the SSW is neither large nor repetitive. The motion of the capsule-toy coming out of the vending machine takes place quietly and only once. The nature of the motion becomes more visible when we examine the concurrent occurrence of the hand gesture. The examination of Kei's hand gestures makes the meaning of the SSW more visible regarding how the capsule-toy comes out.

(3.8) [OSK.4130]

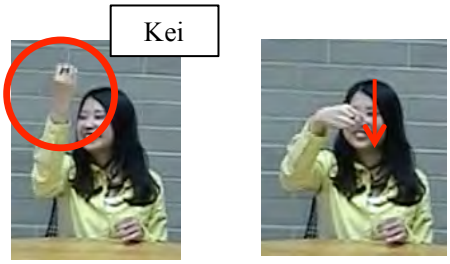


Figure 3.9 *Figure 3.10*

* *

17 Kei: → *atattara pon te,*
win.if SSW QT
'if you win, like *pon*'

While Kei vocalizes *atattara* 'if you win,' she prepares for the subsequent hand gesture by making a soft fist with her right hand, and bringing it to her head level (Figure 3.9). When the subsequent SSW *pon* is vocalized, she swiftly moves the fingers downward, with the wrist still in the same position (Figure 3.10). The hand gesture is iconic in that the motion depicted by *pon* is the falling motion that only occurs once. Having deployed the hand gesture, Kei makes the meaning visible, enacting how the capsule-toy comes out of the machine.

This section has demonstrated that each participant draws on the flexibility of sound vocalization when acting out the referent through the vocalization of the SSWs, during which the characteristics of vocal gestures appear in interaction. Moreover, it has shown that the vocal representations of motions are rhythmically coordinated with the speakers' body movements. The concerted occurrence of vocalization and body movements depicts the motions at the physical site of interaction.

3.5 Invisible concepts

This section examines three cases of *pon*-containing SSWs, all of which depict invisible concepts. The section first examines two cases of *pon*-containing SSWs (i.e., *ponpon* and *po::n*), both of which are related to the manner of speech. In the first case, the SSW *ponpon* is followed by the verb of *iu* 'to say.' The combination, *ponpon iu*, is registered as an idiomatic expression in dictionaries. The analysis of the second case (i.e., *po::n*) shows that the speaker phonetically designs the production of the SSW by lengthening the vowel /o/. The third case contains an SSW (i.e., >*ponponponpon*<) that depicts the sound of the Japanese language, which is also an invisible concept. These cases, except for the first case that involves the conventional SSW, show that by manipulating the vocalization of the SSWs, speakers vocally act out invisible concepts. The analyses of all the cases reveal that, despite the fact that the concept is invisible, the concurrent hand movements visualize the manner of speech or the characteristic of the sound of the Japanese language.

The following segment comes from a talk among three friends who used to take the same course in college. In the following segment, Moko is giving advice to a participant who has been worried that her opinion is not appreciated in a class in which she is currently enrolled.

(3.9) [SMM.1514(ad)] (overlapped turns are not included)

10 Moko: → *kinisezu ponpon yutte ii to omou yo,*
worry.NG SSW say.and good QT think FP
'I think you can say something like *ponpon* without worrying what others will think'

The SSW *ponpon* is registered as a word and its meaning is defined in dictionaries. Specifically, when it is followed by the verb *iu* 'to say' forming *ponpon iu*, it describes the manner of "saying something without due consideration" (Kakehi, Tamori & Schourup, 1996b:996). A noticeable fact is that the vocalization *ponpon* is not prosodically marked (i.e., no pitch change, no sound lengthening). The absence of prosodic marking may be due to the fact that this particular use of *ponpon* is registered in dictionaries. The meaning has been understood as shared among the participants and there is no further need of pinning down the meaning through iconic vocalization. At the same time, vocalizing the SSW without changing the prosody indicates that the SSW is meant to carry its conventional meaning.

Although the meaning of *ponpon iu* 'to say something like *ponpon*' is explained in dictionaries, hand gestures start taking place while Moko vocalizes the SSW. The following Figures in (3.12) show how Moko's hand gestures depict the way in which words come out of her mouth. The hand shape and movements (i.e., opening fingers right in front of her mouth) correspond with the meaning of *ponpon iu* defined in dictionaries, while its production timing is not synchronized with vocalization of the SSW.

(3.10) [SMM.1514(ad)]

Moko


Figure 3.11


Figure 3.12


Figure 3.13

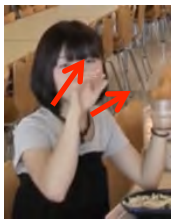


Figure 3.14

* * * *

10 Moko: → *kinisezu ponpon yutte ii to omou yo,*
worry.NG SSW say.and good QT think FP
'I think it's okay to say something like *ponpon* without worrying what others will think'

In line 10, while producing the initial word *kinisezu* 'without worrying,' Moko starts preparing her hand gesture by bringing up her right hand up to the level of her mouth. When she pronounces the second *po* in *ponpon*, she sets her hand at the starting position before the stroke (Figure 3.11). The stroke takes place when she vocalizes the initial sound of the word *yutte* 'say and' (Figure 3.12). Moko spreads her fingers in front of her face. She makes another version of the same hand shape with smaller movements while uttering *ii to omou* 'I think it's okay' (Figures 3.13 and 3.14). In short, the stroke takes place twice, but the occurrence of the strokes is not timely synchronized with the sound beat of the SSW *ponpon*, which contains two occurrences of *pon*. Again, the *ponpon iu* 'say like *ponpon*' is a registered idiomatic expression. It is likely that all the participants know the meaning, and there is no need of pinning it down by means of specific hand gestures. However, when we closely take a look at how the turn unfolds, what *ponpon* refers to has not been clearly vocalized by the speaker at the moment of SSW production. More precisely, what *ponpon* modifies will not be publicly available until Moko

produces the following verb *iu* 'to say.' This case shows that gesture synchronization is related to how the turn unfolds.

By the end of line 10, Moko's turn unfolds in the form of an advice: *kinisezu ponpon yutte ii to omou yo* 'I think you can say something like *ponpon* without worrying what others will think.' Immediately after, she produces another SSW *po::n*, followed by the quotative particle *te*, as an increment in line 11. By virtue of its placement right after the advice on how to speak, the SSW *po::n* is also understood to articulate the manner of speech in relation to the advice.

(3.11) [SMM.1514(ad)]

10 Moko: → *kinisezu ponpon yutte ii to omou yo,*
worry.NG SSW say.and good QT think FP
'I think you can say like *ponpon* without worrying what others will think'

11 → *po::n te*
SSW QT
'like *po::n*'

The SSW *po::n* is a variation of *pon*, in which the vowel /o/ is lengthened. Moko's phonetic design of the SSW indicates that how to say something takes longer in some way. As stated in the Introduction of this chapter, *pon* is defined by dictionaries in many ways, one of which matches how Moko designs her turn by deploying hand gestures. The definition of *pon*[13] is "the manner of hitting or kicking or throwing something lightly or with momentum" (Ono, 2007:433). By lengthening the duration of the vocal production into *po::n*, Moko is likely to act out the word thrown away into the air. Naturally, when you throw something in the air, it will land on the ground shortly after. Therefore, the SSW *po::n* vocally acts out the manner of speech, in particular, the process by which words come out of the mouth.

[13] The original Japanese wording is: *karugaru to ikioi yoku uttari kettari hanattari suru sama.*

Although the dictionary definition does not precisely refer to the manner of speech, the definition is consistent with how Moko designs the concurrent deployment of hand gestures as well. Upon the vocalization of the lengthened version *po::n*, Moko makes the same hand movement she has previously made (Figures 3.11 and 3.12), which is spreading fingers in front of her mouth. When she previously produced the hand movement, we observed that the hand movement was not synchronized with her speech. In contrast, this time Moko's hand movement is well synchronized with the sound beat, thereby visually depicting the way she throws words out of her mouth. The following transcript along with Figures illustrate the movement and its timing.

(3.12) [SMM.1514(ad)]

10 Moko: → *kinisezu ponpon yutte ii to omou yo,*
worry.NG SSW say.and good QT think FP
'I think you can say like *ponpon* without worrying what others will think'

Figure 3.15 *Figure 3.16*

**

11 → *po::n te*
SSW QT
'like *po:n*'

Immediately before the production of the SSW *po::n,* Moko's hand is positioned in front of her mouth. As she vocalizes *po::n*, the hand quickly goes forward while opening her fingers (Figures

3.15 and 3.16). This hand movement unfolds in unison with the beat of *po::n.* By coordinating the vocalization of the SSW and her concurrent gestures, Moko designs the production of the SSW as if she were throwing words out of her mouth into the air. Thus, Moko's hand gesture visually and iconically depicts what the SSW referred to. It should be noted that, in Figure 3.14, Moko directs her gaze at her right side, to a stranger who is about to pass by accidentally.

The current segment contains the SSW *po::n*, also seen in segment 3.4 (basketball). Let us compare the two occurrences. In the basketball case, the speaker first deployed a version of the SSW *pon*, depicting the action of dribbling the ball, which took place close to the floor, in a low position. Then the speaker enacted the player's shooting movements using the SSW *po::n* in a high pitch. That is, the higher-pitched vocalization sound-iconically differentiated the shooting action from the preceding dribbling action. In contrast, the occurrence of *po::n* in the current segment, which does not involve pitch change, is idiomatic, and used in a figurative sense. Furthermore, the referent of *po::n* is not associated with an noticeably high or low position. In short, the pitch design associated with SSW vocalization is unique to what the speaker depicts in a specific environment during face-to-face interaction.

The next segment comes from the same conversation, approximately 34 minutes later. The SSW appearing in this segment also concerns an invisible concept, which is the sound of the Japanese language. It should be noted that the participants took the same seminar, in which they studied second language acquisition theory, and they have knowledge of linguistics as well. Prior to this segment, the participants have been discussing their common Japanese friend, whose English sounds *mairudo* 'mild' and *marui kanji* 'roundish,' and how such vocal quality conveys the impression that she is a proficient speaker of English. The following segment begins with *ato sa* 'and' (line 1), which indicates that Ayaka is going to provide an additional explanation

presumably related to either the common friend or the characteristics of speech sound. The SSW >*ponponponpon*< appears in line 5.

(3.13) [SMM.4923(ad)]

1 Ayaka: *=ato sa,*
and FP
'and'

2 *nihongo tte sa*
Japanese QT FP
'as for Japanese'

Figure 3.17

Figure 3.18
(Enlarged version of 3.17)

*

3 *boin ga goko shika nai ka: sa:*
vowel SB five only nonsexist thus FP
'there are only five vowels, so'

4 Moko: *nn*
'uh uhu'

5 Ayaka: → *nanka >ponponponpon< tte [saa*
like SSW QT FP
'like >*ponponponpon*<'

6 Moko: *[un*
'uh huh'

7 Ayaka: *kicchiri wakareteru*
exactly separated.ing
'the sounds are all separated'

In line 2, Ayaka raises the topic of the Japanese language by mentioning *nihongo tte sa* 'as for Japanese.' She then introduces *boin* 'vowels' (line 3), marked by the subject marker *ga*. While uttering *boin ga goko shika nai ka: sa:* 'there are only five vowels, so,' Ayaka produces hand gestures, spreading the fingers of her right hand, the palm facing toward the recipients sitting in front of her (Figures 3.17 and 3.18). Subsequently, using the fingers of her left hand, she draws a half arc with the tips of her right-hand fingers. It is important to note that the number of fingers is five, which is the same as the number of Japanese vowels. By making this drawing movement, she visualizes the Japanese sound system, which consists of 'only five vowels.' This unit is marked by the conjunctive particle *kara*, vocalized as *ka::*, 'so' or 'thus,' indicating that the subsequent element is related to the Japanese language and its five vowels.

At this moment, Ayaka vocalizes the SSW >*ponponponpon*<. By virtue of its placement in this sequential position, this SSW is heard as a depictive element articulating how the Japanese language sounds. The SSW >*ponponponpon*< consists of four occurrences of *pon,* and is vocalized more quickly than the surrounding talk. The number four corresponds to the number of the divides of five fingers that have been used as an analogy. The fast vocalization of the SSW iconically acts out how Ayaka hears the speech sounds of Japanese. By manipulating the vocalization of the SSW, Ayaka vocally acts out the impression of the sound of the Japanese language.

This SSW >*ponponponpon*< is followed by the quotative particle *te*, which is then followed by *kicchiri wakareteru* 'the sounds are all separated.' These elements that follow the SSW indicate that the SSW >*ponponponpon*< depicted how the sounds of the Japanese language are separated. Here, a piece of background information regarding the Japanese sound system can be useful for the reader to understand Ayaka's production of the SSW. Unlike English, the

rhythmic unit of Japanese is called mora, which contains one vowel at the end. Furthermore, as Ayaka mentioned, Japanese has only five vowels (i.e., /a/, /i/, /u/, /e/ and /o/). Her use of the *pon*-containing word does not correspond to any of the dictionary-meanings of *pon* outlined in Section 3.3. In this segment, *pon* is deployed to depict how the Japanese language sounds separated.

A close examination of concurrent hand gestures supports this claim that the SSW depicts how Japanese language sounds separated.

(3.14) [SMM.4923(ad)]

Figure 3.19

*

3 Ayaka: *boin ga goko shika nai ka: sa:*
vowel SB five only nonsexist thus FP
'there are only five vowels, so'

4 Moko: *nn*
'uh uhu'


Figure 3.20

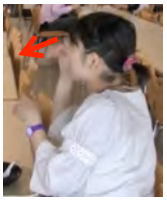


Figure 3.21

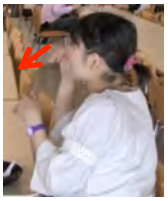


Figure 3.22


Figure 3.23

* * * *

5 Ayaka: → *nanka >ponponponpon< tte [saa*
like SSW QT FP
'like >*ponponponpon*<'

Immediately before the production of the SSW, Ayaka initiates the preparation of her hand gesture. At the end of line 3, while producing the final particle *sa*:, Ayaka sets the right hand at her eye level. She bents the fingers at a ninety-degree angle, and places her fingers parallel to the table (Figure 3.19). The subsequent production of the SSW is rhythmically synchronized with her hand strokes. When vocalizing >*ponponponpon*<, Ayaka mimics the rhythms with her hand four times, while the tips of her fingers go downward step by step, as if her hand was marking four steps of a ladder (Figures 3.20 to 3.23). That is, although the sound of Japanese constitutes an invisible concept, the speaker visually designs the production of the SSW by means of hand gestures that occurs along with vocal representation. By coordinating vocal and visual gestures, Ayaka embodies how she hears the sounds of Japanese and their separations.

This section has examined three cases of SSWs that depict invisible concepts. The speakers vocally act out invisible concepts by making use of the flexibility of vocalization, for example, by changing the rate and duration of the SSW as well as by repeating *pon*. Moreover, concurrent production of hand gestures makes speech content visible to the recipients.

3.6 Motionless objects

This section examines two cases in which SSWs are deployed in depicting arrangements of motionless objects (i.e., <*ponponponpon*> and ↑>*ponponpon*<) that are absent from the actual site of their interaction. By repeating *pon* several times and manipulating the rate and pitch of speech, speakers phonetically design the vocalization of the SSWs, thereby vocally acting out the existence of motionless objects and the number of existing objects. Moreover, while vocalized

SSWs represent objects arranged in a certain space, the concurrent hand gestures visually depict how the objects are arranged.

The first segment comes from a talk among friends at an English language school in the United States. The participants talk about writing assessments at school. Prior to the segment, Shun asked Miki: *nani midashi tte* 'what do you mean by headings?' In the segment below, Miki answers to the question by explaining the page layout of the English textbook, which the participants presumably use at school. By deploying the SSW, Miki acts out what she has meant by *midashi* 'headings.' The SSW that appears in this segment is <*ponponponpon*>, in which four occurrences of *pon* take place in a row.

(3.15) [DNS.0122]

Figure 3.24

*

20 Miki: *mini topikku ga koo atte saa,*
mini topic SB like.this exist.and FP
'there is a mini topic like this'

21 → *<ponpon[ponpon> tte*
SSW QT
'the headings are like <*ponponponpon*>'

22 Shun: *[ah s: so yu koto ne?*
oh so say thing FP
'oh that's what you meant'

In line 20, Miki starts describing the page layout of the textbook. While saying *mini topikku ga koo atte sa* 'there is a mini topic like this,' Miki forms a circle a short distance away from her body, using both hands (Figure 3.24). One of the elements in her turn is *koo* 'like this,' which co-occurs with gestures of some sort that draw attention from the recipient (Streeck, 1993; 2002). Miki's hands are separated and there is a space between them. Her design of this hand gesture indicates that the mini topic is printed at the top of the page. This utterance ends with the continuing form of the verb (i.e., *atte*) followed by the final particle *sa*, indicating that her explanation of the page layout continues. While she produces the hand gesture described above, Miki gazes at Shun, who asked about *midashi* 'headings.' Once Miki secures Shun's gaze, Miki looks down and vocalizes the SSW <*ponponponpon*> more slowly than the surrounding talk. By deploying the SSW <*ponponponpon*> in this sequential position, Miki indicates that the SSW depicts the location of *midashi* 'headings' in question in relation to what Miki calls 'mini topic,' which is printed at the top of the page.

The SSW consists of four occurrences of *pon*. The number of times the sound is repeated relates to the content of the speech (Hamano, 1998). In particular, Miki acts out the existence of exactly four *midashi* 'headings' by vocalizing the SSW. Importantly, halfway through the production of the SSW, Shun claims his understanding as to what Miki has meant by *midashi* 'headings' in line 22.

Significantly, as the following Figures show, the vocalization of the SSW is synchronized with hand movements, which visually depict speech content. By coordinating speech and hand, Miki not only vocally acts out the existence of the four headings, but further depicts their arrangement.

(3.16) [DNS.0122]

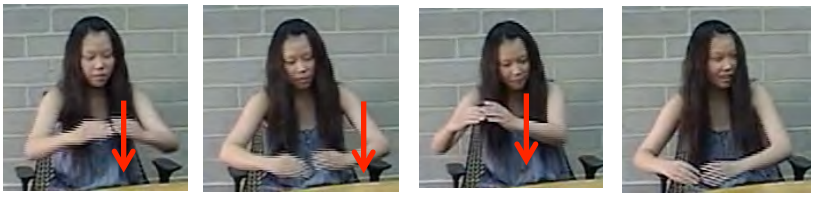


Figure 3.25 *Figure 3.26* *Figure 3.27* *Figure 3.28*

* * * *

21 Miki: → <*ponpon*[*ponpon*> *tte*
SSW QT
'the headings are like <*ponponponpon*>'

Miki forms a circle that is smaller than the last circle she formed. While keeping the same hand shape, Miki brings the hands down and then up as she vocalizes each sound of *pon.* The vocalization accompanied by the hand movement is repeated four times (Figures 3.25 to 3.28). It should be noted that the positions in which she brings her hands down look like each corner of the page below the 'mini topic,' which she has designed as printed in the top of the page. Furthermore, as mentioned, the SSW is vocalized more slowly than the surrounding talk. The slow vocalization may be due to the accurate co-production of the hand gesture while vocalizing each *pon*; she had to move hands aiming at the four corners of the imaginary page. That is, by repeating *pon* four times along with the hand movement, Miki depicts how *midashi* 'headings' are printed on the page.

The next segment also includes a SSW that depicts an arrangement of objects with the visual aid of hand gestures. In the following segment, Rena is talking about her experience of seeing old houses, while driving in a rural area of the United States. She characterizes such areas

as *nanimo nai* ‘there is nothing there’ (line 22). She deploys the SSW >*ponponpon*< in line 23 in acting out how she observed old houses in the area.

(3.17) [OTR.3116]

22 Rena: *su-, nan daro, nanimo nai n da kedo:,*
what wonder nothing nonsexist N CP but
‘like, there’s nothing there but’

23 → *↑>ponponpon< tte furui ie ga ko:,*
SSW QT old house SB like.this
‘old houses are like >*ponponpon*<, like this’

24 *mo: daremo tsukattenai no kanaa tte yuu*
already nobody use.and.NG N wonder QT say
‘like nobody seems to use them’

25 *furui ie aru jan [tamani:,*
old house exist TAG sometimes
‘we see such old houses sometimes, huh?’

In line 22, Rena describes the scene of a particular rural area she has seen while driving: *nan daro, nanimo nai n da kedo:*, ‘like, there’s nothing there but.’ The final *kedo* ‘but’ indicates that the subsequent element is something contrastive to what she has seen. That is, the subsequent element will be relevant to the view she has seen in the rural area in which there was nothing. At this very moment, Rena produces the SSW ↑>*ponponpon*<, which consists of three occurrences of *pon.* As indicated in the transcript notation, Rena manipulates the vocalization of the sound by repeating *pon* three times as a word, vocalizing it more quickly than the surrounding talk. Furthermore, she vocalizes it in a higher pitch. By designing the turn this way, she vocally acts out something that exists in the place in which there was nothing. Immediately after, Rena continues as *furui ie ga ko:* ‘old houses are, like this’ (line 23). Rena’s turn makes explicit that the deployment of the SSW acts out how old houses are located in the area. It might

be the case that the production of >*ponponpon*< at a higher speed rate iconically corresponds how she has observed the houses from the car window while driving.

The SSW co-occurs with hand gestures. However, unlike other cases of creative uses of the SSW *pon* analyzed thus far in this chapter, the hand movements are not synchronized with the sound beat. The following is the extracted transcript along with Figures of significant moments of the turn design.

(3.18) [OTR.3116]

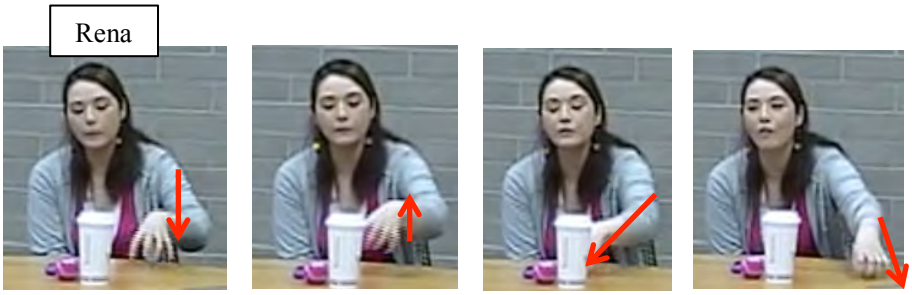


Figure 3.29 *Figure 3.30* *Figure 3.31* *Figure 3.32*

* * * *

23 Rena: → ↑>*ponponpon*< *tte furui ie ga ko:,*
SSW QT old house SB like.this
'old houses are like >*ponponpon*<, like this'

The Figures 3.29 to 3.32 show how Rena moves her hands. While vocalizing >*ponponpon*<, she places the tips of her left-hand fingers on the table (Figure 3.29), and pushes them lightly toward the table. While producing the following quotative particle *tte*, she moves up her hand (Figure 3.30). From this point onward, she draws a zigzag movement from the front of her body toward the front right (Figure 3.31), and then toward the front left, further away from her body (Figure 3.32). At this moment, Rena vocalizes *ko*: 'like this' that has been found to co-occur with some

sort of hand gestures that draw attention from the recipient (Streeck, 1993, 2002). At each point of the zigzag figure, she lightly pushes the fingertips onto the table. As presented in the above transcript, which includes Figures, this zigzag hand movement is not synchronized with the sound beat of the SSW. Here, let us note that what > *ponponpon* < refers to is not verbalized yet. Rena makes it clear right after the SSW followed by the quotative particle *tte*: *furui ie* 'old houses.' Although I expressed the movement as zigzag, what is important is that the corners of the zigzag points are all separated. By designing the hand movement as a zigzag, Rena recreates the original scene of viewing, in which houses show up sporadically.

The two segments examined in this section reveal that speakers utilizes sound flexibility, while acting out the existence of the referent by vocalizing the SSWs. In particular, sound repetition was the major practice when acting out the motionless objects arranged in/across a certain space. Both SSWs were not related to dictionary definitions, but depicted the existence of objects along with the concurrent hand gestures.

With respect to the hand gestures, the first case (textbook) was synchronized with the hand gesture. In contrast, the second case (old houses) was not in sync with the vocalization of the SSW. The crucial difference between the two cases was how the turn that incorporates the SSW emerged in the interaction. In the first segment, Shuu asked Miki what *midashi* 'headings' was, which opened up the insert sequence. Thus, the referent of the SSW description had already been clarified before Miki started the turn that incorporated the SSW <*ponponponpon*>. Differently put, Miki was responsible for providing the depiction so that the recipient understands what *midashi* 'headings' exactly referred to, which in turn, involved closing the insert sequence. On the contrary, in the second segment, Rena was engaging in the activity of telling, and the referent of the SSW description had not yet appeared at the moment of the SSW

deployment. The recipients had no access to what was ↑>*ponponpon*<. That is, whether or not the referent is clear at the moment of vocalizing the SSW seems to determine synchrony of gesture and speech.

3.7 Summary of chapter 3

This chapter examined eight cases of *pon*-containing SSWs, demonstrating how SSWs operate as vocal gestures in face-to-face interaction. The analyses showed that even the single *pon* sound can be phonetically designed in various ways. In each case, the speaker flexibly manipulated the vocal production of the SSW *pon* and its variations, for example, by changing the rate and pitch. Such vocalization was deployed while the turn unfolds, thereby vocally acting out each respective referent. The number of times the morphological unit *pon* was repeated was also one of the factors in determining what each SSW depicts.

The analyses also showed that the vocalization of SSWs was accompanied by hand gestures. By deploying visible hand gestures, in addition to vocally acting out the referent through the SSW, speakers visually depicted various types of referents at the current site of the interaction: motions, invisible concepts, and arrangements of motionless objects. That is, various multimodal semiotic resources were all coordinated in depicting the referent through the deployment of the SSWs in interaction.

The chapter primarily concerned SSWs as vocal gestures; however, the analyses also clarified notable phenomenon around hand gesture occurrences upon the deployment of SSWs. As noted, Kita (1993, 1997) confirmed that adverbial SSWs, in particular motion adverbial SSWs, are almost always accompanied by hand gestures. This chapter provides support for his finding, while expanding it further: The SSWs that depict different types of referents—

motionless objects and invisible concepts—can also be accompanied by hand gestures. Also, it is important to emphasize that not all the hand gestures rhythmically synchronized with the beat of the SSW as we saw in the case of the idiomatic expression *ponpon iu* 'to say like *ponpon*' and in the case of old houses. Further, it does not necessarily mean that all the conventional SSWs do not rhythmically synchronize with gestures. The case of capsule-toy was one example. The use of *pon* in that example fit the dictionary definition, but it was accompanied by a timely synchronized hand gesture. That is, the relationship between vocalization of SSWs and gesture occurrence cannot simply be explained in terms of the degree of conventionality of SSWs. Rather, how turns that incorporate SSWs unfold appeared to be a more crucial agenda. What we have to care about is whether or not the target of the SSW description is clear to all the participants upon the deployment of the SSW. A further systematic analysis will clarify when and how gestures co-occur upon and synchronize with the vocalization of SSWs.

Part 3

Analysis:

What Speakers Can Do through Deployment of Sound Symbolic Words in Talk-in-interaction

Chapter 4

Concretizing Absent Objects through Deployment of Sound Symbolic Words

4.1 Introduction

This chapter examines vocalization of sound symbolic words (SSWs) that concretizes objects absent from the interactional space. In the current literature, there have been different understandings of what objects are. For example, Navile et al. (2014a) consider objects "those elements of the physical world that we can experience sensorially, i.e. that we typically see, hear and touch" (p. 5). On the contrary, other scholars deal with immaterial, intangible, and/or imaginary objects (e.g., Keevallik, 2014). This chapter centers on absent objects, in particular, with objects that are not present in the interactional space but exist as physical objects in the real world. In other words, it examines objects in interaction that are created through the speaker's memory and imagination. As Matthews (2014) stated, "there is in principle no difference between unpacking the meaning and consequence of objects on account of their perceptual and physical properties" (pp. 385-86). This chapter uses the phrase "properties of objects" to refer to "whatever is observable, noticeable, describable, experiential about them" (Matthews, 2014:384). In Japanese, the deployment of SSWs is key to understanding how properties of objects appear in interaction. As Chapter 2 (SSWs and Bodily Conduct) and Chapter 3 (SSWs as Vocal Gestures) have demonstrated, the deployment of SSWs vocally gesticulates and often accompanies hand gestures. Through a number of concurrent semiotic resources, SSWs are capable of instantly depicting the properties of objects.

There are two main streams in our understanding of objects in interaction (Nevile et al., 2014a). On the one hand, participants may use physical objects as situated resources while the

ongoing interaction unfolds. For example, situated objects can operate when participants take turns (Day & Wagner, 2014), shift from one activity to another (Nielsen, 2014; Richardson & Stokoe, 2014), or change the participation framework (Mikkola & Lehtinen, 2014). On the other hand, participants may use objects for practical accomplishments such as cooking (Mondada, 2014), crafting (Ekström & Lindwall, 2014), and dance instruction (Keevallik, 2014). In order to accomplish certain practical goals, participants create or reshape objects in a way that enables the participants to understand the created and/or reshaped objects.

However, Sakai et al. (2014) claimed that these two ideas cannot be treated separately, because "[f]or objects to be used as interactional resources, they must first be made recognizable and intelligible" (p. 339). This chapter is aligned with this claim, and demonstrates that SSWs depict objects that serve as situated resources, that is, objects that help participants to achieve interactional accomplishments, such as making suggestions and searching for a word (e.g., C. Goodwin & M. Goodwin, 1986; M. Goodwin, 1983; Hayashi, 2003b; Lerner, 1996).

This chapter is organized as follows. Section 4.2 demonstrates how speakers concretize an absent object in the interactional space via verbal and visual conduct. Speakers first draw the object by referring to the target object with the medial demonstrative pronoun *sore* 'that one.' Then, the speakers continue the turn, during which SSWs depict the property of the object, thereby accomplishing an interactional achievement, in particular seeking confirmation and making suggestions. Section 4.3 reveals that depiction via SSWs is not always satisfactory; sometimes it has limitations. The analyses show that the speaker occasionally needs to reformulate what has been depicted through SSWs with non-sound symbolic words (non-SSWs) in order to achieve intersubjectivity. This study refers to formulation as "the production of a

particular description, categorization, representation, or lexical selection" (Prior, 2016:208)[1] and reformulation as "subsequent descriptions, categorizations, summaries, or revisions of prior formulations" (ibid).[2] Section 4.4 examines cases in which SSWs operate as a provisional element in word search activities. It first outlines previous research on word search activities. The following analyses demonstrate that when the speaker is searching for a word, the practice of concretizing an absent object serves as a way of referring to the target objects without verbalizing its name. Section 4.5 concludes this chapter by summarizing its findings.

4.2 Instant depiction of absent objects

This section demonstrates that speakers concretize absent objects in face-to-face interaction in a particular way in which the recipients can understand these invoked objects. It aims to reveal how turns that incorporate SSWs unfold, during which the absent object(s) is/are drawn into the interaction and their properties depicted. As described in previous studies, the speaker typically refers to the name of an object or refers to an object that has been talked about using a *so*-series, or 'that-series,' demonstrative pronoun. *So*-series demonstratives involved in this chapter are *sore* 'that one' and *sono* 'that' that is always followed by a noun (e.g., *sono pen* 'that pen'). By deploying one of these demonstrative pronouns, the speaker establishes the relevance of the object in the ongoing interaction. The subsequent vocalization of SSWs then instantly depicts the properties of objects: for example, how the object works.

The first segment comes from a conversation among three friends: Moko, Yasu, and Ayaka. Prior to the segment below, Moko told a surprising story about her past experience: when Moko exited the gate of the Akihabara station, she saw a group of people dressed up as anime

[1] Prior's definition is based on previous studies such as Antaki (2008), Antaki and Widdicombe (1998b), and Bilmes (2011).
[2] Prior integrates the idea of upshot and gist formulations introduced in Heritage and Watson (1979).

characters. The teller Moko was working at a shop in Akihabara, which is famous as a shopping area for electronic appliances, computers, and anime-related products. According to Moko, there was a promotional event for releasing a new model of mobile phone designed for big fans of the anime series *Evangelion*. Moko assessed the Akihabara district as *sugoi yo* 'it's amazing' and the two recipients Yasu and Ayaka also assessed Akihabara as *dokutoku da yo ne* 'it's unique, isn't it?.' After a silence, the following segment starts. This is when the story recipient Yasu opens up a new sequence in which they talk about the costume in question. In the segment, the SSW is first deployed by Yasu who seeks confirmation of the property of the object, *suutsu* 'costume,' mentioned in the telling. By doing so, Yasu claims that she shares the knowledge of the object with the teller. It is worth noting that this SSW is a creative use of the conventional SSW *pitchiri*, which is defined as "*sukima naku micchaku shite iru sama* 'the manner of things stuck together without gaps'" (Ono, 2007:370). Responding to Yasu's claim, the teller Moko confirms the quality of the *suutsu* 'costume' being *pitchiri*. That is, two different speakers deploy variations of the SSW *pitchiri*, while they depict the shape of the costume through juxtaposed verbal, prosodic, and/or bodily resources.

(4.1) [SSM.010017]

Figure 4.1

*

1 Yasu: .h(.) te(h)ka so↑no suutsu sa:a,
uhm that suit FP
'uhm, that costume is'

Figure 4.2

Figure 4.3

2 → *pit↑chiri shite nai? kek[koo.]*
SSW doing FP fairly
'being *pit↑chiri*, huh? Fairly'

Figure 4.4

Figure 4.5

3 Moko: → *[ah pit]tchiri shiteru.*
oh SSW doing
'oh, that's being *pittchiri*'

4 Yasu: *hhh*

In line 1, Yasu, the story recipient, draws the object into the interaction by referring to the name of the object, preceded by the demonstrative *sono* 'that,' as in *sono suutsu saa* 'that costume is.' Yasu's turn makes the object—costume—relevant to the forthcoming talk. When Yasu produces the final particle *sa*, she places her palms facing her chest, and maintains this position for a moment (Figure 4.1).

Yasu then vocalizes the SSW *pit↑chiri*, in which the second half *chiri* is emphasized and produced in a high pitch. What follows is *shitenai?* 'huh?,' through which Yasu seeks Moko's confirmation of the property of the costume. What is visibly noticeable is that while Yasu

vocalizes *pit↑chiri shitenai?* 'being *pit↑chiri*, huh?,' she draws an S shape with her hands along the line of her body (Figures 4.2 and 4.3). Through the coordination of various multimodal resources, such as hand, body and voice as interactional resources, Yasu adds the property of the object, costume (more precisely, people wearing the costume), thereby concretizing it in the interactional space. As mentioned, dictionaries define *pitchiri* as "*sukima naku micchaku shite iru sama* 'the manner of things stuck together without gaps'" (Ono, 2007:370). This definition does not specify how the costume is stuck together with her body, but the analysis of the actual interaction reveals that the concurrent hand movements visually depict an aspect of the property that has not been verbally produced. Before Yasu completes her turn with the adverb *kekkoo* 'fairly,' Moko produces the change-of-state token *ah* 'oh' (Heritage, 1984) and then recycles the expression *pitchiri suru* that Yasu produced when seeking confirmation. During the production of this turn, Moko touches her chest with her hands (Figures 4.4 and 4.5). It is remarkable that Moko's version of *pitchiri* slightly differs from Yasu's. Moko takes some time before producing *chi*, as transcribed by the colon in *pit:chiri*, and emphasizes the second half *chiri*. It might be the case that Moko, who has firsthand knowledge of the costume, uses this vocal design to upgrade the quality of tightness.

This segment has demonstrated that the speaker drew an object into the ongoing interaction which had been mentioned in the prior talk. The *so*-series 'that-series' demonstrative, in this particular case, *sono suutsu* 'that costume' was at work. The speaker then depicted the property of the object through deployment of the SSW, which was accompanied by vocal and body gestures. The next case also shows the same practice of concretizing objects.

The participants in the following segment are graduate teaching assistants: Tama, Mie, and Nao, who teach a class on a daily basis. The recording was conducted in their office. The

topic at hand was Tama's *yubisashiboo* 'pointer stick' which Tama, who is sitting on the left, holds on her lap (Figure 4.6). The talk then moved on to the topic of overhead projectors, which the three participants use in their everyday teaching. Tama and Mie agreed that the light of the overhead projector is so bright that it makes them feel dazzled. Immediately after closing this sequence, Mie opens a new sequence in which she suggests a hypothetical function of *yubisashiboo* 'pointer stick.' The analysis focuses on how Mie makes the object relevant to the subsequent talk, and depicts the property of the object by deploying the SSW. In line 61, the first line of the segment, the conversation about overhead projectors is coming to the end.

(4.2) [HKK.1512]

61 Mie: ˚˚*sokka, iyaiya*˚˚
I.see well
'˚˚ I see, well˚˚'

Figure 4.6

*

62 *nanka sore sa,*
like that.one
'like, that thing'

Figure 4.7

Figure 4.8

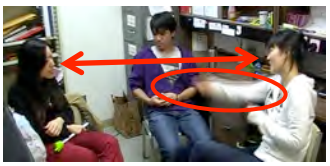


Figure 4.9

Figure 4.10

* * * *

63 → *botan toka oshite >hyun< te nobiru [yatsu toka*
button etc. push.and SSW QT extend one etc.
'how about if it extends like ↑ >*hyun*< by pushing something like a button'

64 Tama: *[ah:*
'yeah'

65 *na- na [ttara ii*
become.if good
'that's gonna be great'

In line 62, by deploying the demonstrative *sore* 'that one,' while gazing at the real thing on Tama's lap, Mie draws the *yubisashiboo* 'pointing stick' into the interaction. When Mie produces the following final particle *sa*, she brings her hands in front of her body, as shown in Figure 4.6. At this moment, her hand starts adopting the position of holding a stick. Figure 4.6 also shows Tama's gaze directed to Mie's hand.

In line 63, while Mie mentions a hypothetical button, *botan toka* 'something like a button,' Mie looks at her own hands, which are shaped as if they were holding a stick (Figure 4.7). Tama's gaze has also been directed to Mie's movements. In Mie's turn, *toka* in *botan toka*

'something like a button' indicates the nature of the hypothetical talk: the content of the talk is not real, thus Mie designs her turn so that she does not definitely say that the thing to push is a button. What follows is *oshite* 'push and,' during which Mie moves the thumb as if she were pushing the button on the stick (Figure 4.8). It is also noticeable that she does not direct her gaze toward her hands now, but has moved forward, staring into the air.

Mie then vocalizes the SSW ↑>*hyun*< , which is produced in a higher pitch and more quickly than the surrounding talk. The vocalization is synchronized with a wide movement of her right arm: Mie quickly starts stretching out her arm forward (Figure 4.9), further extending it upward (Figure 4.10). By performing this movement in sync with the vocalization, Mie acts out the way the stick transforms its shape after pushing the hypothetical button. Mie's verbal and visual conduct constitute the object within the interaction. While Mie makes these hand and arm movements, she holds a mutual gaze with Tama (Figures 4.9 and 4.10). The following grammatical element again involves *toka*, as in *yatsu toka* 'something like,' contributing to making the content of the utterance indefinite.

It is worth noting that, as soon as Tama hears Mie's vocalization of the SSW (line 63), Tama comes in by producing the acknowledgement token *ah:* 'yeah' (line 64). Tama then continues her turn by appreciating the hypothetical function Mie has suggested as *nattara ii* 'that's gonna be great.'

In this segment, the SSW was deployed to depict how the object transforms its shape hypothetically. In particular, the transformation process is depicted in a moment that involves rapid movements of the speaker's arm and hand. The object was drawn into the talk initially by the speaker's use of the demonstrative *sore* 'that one,' which referred to the object that had been talked about.

This section has demonstrated how the speaker draws the object into the ongoing interaction and depicts the property through the deployment of SSWs. When drawing the object, the speaker utilized one of the *so*-series 'that-series' demonstratives and made the object mentioned in the prior talk relevant to the ongoing interaction. The drawn object then acquired its properties when the speaker deployed the SSWs. Upon the vocalization of the SSW, the juxtaposed multimodal resources contributed to concretizing objects that did not exist in the interactional space. The analyses have also shown that the concretized objects served as an integral part of the interactional accomplishments that the speaker had engaged in.

4.3 Possibilities and limitations of SSWs

In interaction, the SSWs' power of depiction entails possibilities and limitations. As introduced in the Introduction, possibilities refer to the SSWs capacity of instant depiction of various referents. In contrast, limitations refer to the restriction of the deployment of SSWs when its sequential environments do not allow the recipient to make sense of the depicted referents. This section demonstrates how participants' conduct mobilizes these two features in interaction. It examines cases in which objects are introduced, while the turn unfolds, through the deployment of the proximal demonstrative pronoun *kore* 'this one,' during which he/she employs iconic hand gestures (McNeill, 1992). In face-to-face interaction, the demonstrative pronoun *kore* 'this one' draws the recipients' attention to the co-occurring hand gestures that make the shape of objects (e.g., Hayashi, 2005a; Streeck, 1993). In this sense, the deployment of *kore* 'this one' is different from that of *sore* 'that one,' which simply refers to objects that have been mentioned earlier. This section examines cases in which the object that the speaker has drawn into the interaction using a *ko*-series 'this-series' demonstrative then gets its properties through the subsequent

vocalization of SSWs. The object referred to with a *ko*-series 'this-series' demonstrative is not always available in the interactional space. The object is usually something available in the real world, but might be unclear to the recipients, because the existence has not been necessarily established among the participants in the interaction at the moment of its production. This fragile status of the object discloses the key dimensions of the deployment of SSWs: possibilities and limitations.

The participants in the next segment are Aiko, Yumi and Rina, who are colleagues at work. Prior to the segment below, Aiko informed the recipients that she had tenosynovitis on her wrist. Yumi and Rina provided possible reasons for Aiko's discomfort as *hatarakisugi* 'working too hard' (by Yumi) and *mausu* 'mouse' (by Rina). Yumi then suggests buying a new mouse in line 9 in (4.3).

In order to illustrate the constant visual display as performed by the participants' hand shapes, the segment is divided into three parts. The first part (4.3) focuses on how the object in question is drawn into the interaction by means of hand gestures. The second part (4.4) shows how the hand shape is kept over time, displaying the continuing existence of the object in the interactional space. The third part (4.5) then demonstrates how the property of the object is audibly and visibly added through the vocalization of the SSW.

Prior to the deployment of the SSW, the speaker produces the demonstrative pronoun *kore* 'this one' that co-occurs with hand gestures depicting a specific object: a mouse pad that has good support. Notwithstanding the visual display of the object that the speaker has rendered up to this point, the speaker audibly adds a property, bouncing quality, of the pad by vocalizing the SSWs. It appears that the speaker counts on the possibility presented by vocal gestural aspect of the SSW on its deployment.

(4.3) [OL.1934-1]

9 Yumi: *kaikae ↑na, °mausu°*
renew FP mouse
'you should buy a new mouse'

10 *tteka katte morai na onezoo-san ni.*
I.mean buy.and receive FP personal.name by
'I mean, ask Onezoo-san to buy it for you'

11 Aiko: [*eh?*
eh
'what?'

12 Rina: [*mausu ga warui no sore wa:*
mouse SB wrong Q that TP
'was it because of the mouse?'

13 Aiko: *iya iya mausu janai,*
no no mouse isn't
'no no it isn't because of the mouse'

14 *mo: kenshooen ya nen mukashi kara*
like tenosynovitis CP FP long.ago since
'like, I've been having chronic tenosynovitis'

Figure 4.11

*

15 Yumi: *ya (nanka) mausu to sa:*
no (like) mouse and FP
'no, (like) a mouse and'

Figure 4.12

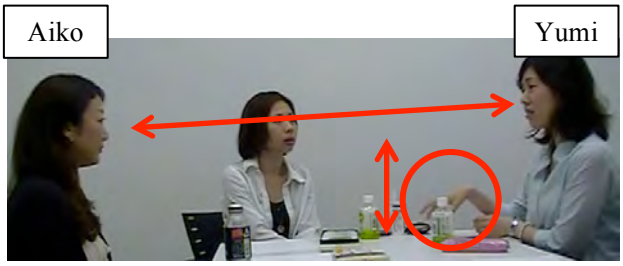


Figure 4.13

*

*

16 *koko no patto toka no ii yatsu o sa:*
this.part LK pad etc. LK good one OB FP
'something like a good pad for this part'

17 Aiko: *[ah: [:*
I.see
'I see'

18 Rina: *[ah: [:*
I.see
'I see'

19 Yumi: *[kat[temorau dake demo*
buy.and.receive only even
'even by getting those items'

20 Aiko: *[so (re ka)na::*
(that.one) (wonder)
'I wonder if that helps'

21 Yumi: *sugoi [raku n naru yo*
very ease to become FP
'your pain will be so relieved'

In line 9, Yumi suggests that Aiko buy a new mouse. Yumi then reframes the suggestion as Aiko asking the boss to buy a new mouse (line 10). Aiko initiates a repair by producing *eh?* 'what?' (line 11). Aiko and Rina then clarify that the cause of the pain in Aiko's wrist is not using a mouse but her chronic tenosynovitis (lines 12 to 14). Yumi, who has suggested the cause being

the incorrect use of the mouse, produces *ya* 'no,' which denies that the use of mouse might have caused the pain. In addition to *mausu* 'mouse' (line 15), Yumi adds the item *koko no patto toka no ii yatsu* 'something like a good pad for this part' to the list of items for Aiko's boss to buy (line 16). While producing the turn, Yumi starts making a shape with her hands: right hand holding an imaginary mouse and left hand supporting the right hand (Figure 4.11). The hand shape then accompanies up-and-down bouncing movements (Figures 4.12 and 4.13). By keeping the bouncing movement, Yumi concretizes the object—a good pad—in the interactional space. The reformulated suggestion convinces both Aiko and Rina, who appreciate the suggestion by producing *ah::* 'I see' respectively (lines 17 and 18). Yumi then justifies the suggestion: *sugoi raku n naru* 'your pain will be so relieved' (line 21).

In the following part, Rina and Aiko make sure they understand the right angle for one's wrists when working with a computer. As the figures in the following transcript show, both Aiko and Rina locally construct the best angle for the wrist by forming hand shapes.

(4.4) [OL.1934-2]

22 Rina: [*so* (.) *nanka*,
yeah like
'yeah, like'

Figure 4.14

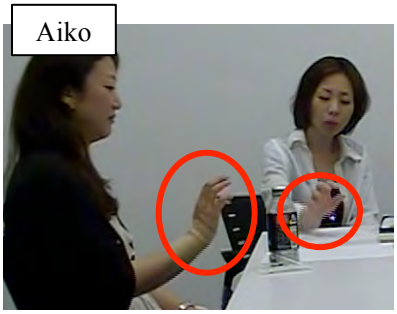


Figure 4.15

23 ↑*ko natten no ga akan nen te na,*
like.this becoming N SB no.good CP QT FP
'doing it like this is no good, it seems'

Figure 4.16

24 *ko: natteruno ee rashii de*
like.this becoming good heard FP
'doing it this way is good, I heard'

Figure 4.17

25 Aiko: *so: na n [ya*
so CP N FP
'is it so'

As these figures show, Rina visually demonstrates a bad angle (Figure 4.14) and Aiko mimics that shape (Figure 4.15). Rina contrasts it against what is presented as a good wrist angle (Figure 4.16). Aiko again mimics the angle with her right hand, and shows appreciation of it: *so: nan ya* 'is that so' (or, in a more colloquial expression, 'I didn't know that'), in line 25 (Figure 4.17).

Then, immediately after Yumi's hands are released from the shape of the good wrist angle, Yumi starts asking Aiko where Aiko's mouse comes from (lines 28 to 32, which are omitted). In the subsequent segment shown in (4.5) below, Yumi comes back offering the same suggestion, in which the SSW *buwan* is deployed. As shown below, Yumi deploys similar hand movements to the ones she has made in (4.3). Moreover, the hand shape that co-occurs with the vocalization of the SSW seems to mimic the angle that has been established as a good angle by Rina and Aiko in (4.4).

(4.5) [OL.1934-3]

33 Rina: [1 ˚*ko:gaku (nan)*˚
optical (CP Q)
'˚ an optical one? ˚'

Figure 4.18

Figure 4.19

34 Yumi: [1 *jaa, kore, kore kattemorai na yo*
then this.one this.one buy.and.receive FP FP
'then, this one, this one, ask your boss to buy it for you'

Figure 4.20

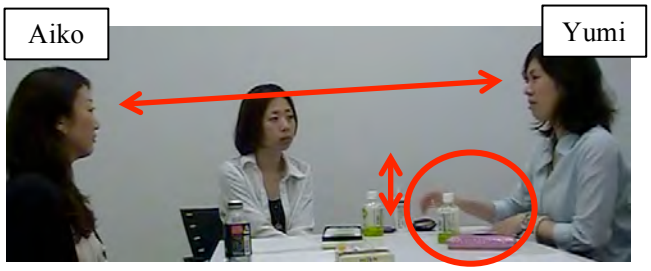


Figure 4.21

35 Yumi: → [2 *buwan to suru yatsu*
SSW QT do one
'the one which goes *buwan*'

36 Aiko: [2 °*so ya naa*°
so CP FP
'°yeah°'

37 Rina?: °° (*kush-*) °°

38 Aiko : °*un wakatta*°
yeah understand
'°yeah, I understand°'

39 [°*katte morau wa*°
buy.and.receive FP
'I'll ask him to buy one for me'

In line 34, by producing *jaa* 'then,' Aiko indicates that the forthcoming element is related to the prior topic (the origins of Aiko's mouse). Here, Yumi verbalizes the demonstrative pronoun *kore* 'this one' twice, while reshaping her hand and reproducing the movements that she has made in line 16 in (4.3). Yumi keeps bouncing her right wrist on her left palm, which is facing upward (Figure 4.18). By performing the movement that goes along with the vocalization of *kore kore* 'this one this one' (line 34), Yumi acts out the property of the object with her body. Yumi then adds *katte morai na yo* 'ask your boss to buy it for you,' which establishes the preceding *kore*

'this one' as the object that Aiko should ask her boss to buy for her. It should be noted that Yumi still keeps the shape and movements (Figure 4.19).

As has been described up to this point, the object—a pad—and its properties seem to have been established. We observed that while the interaction moved forward, Yumi gradually delivered the property of the object via her hand shape and its movements. However, Aiko, the recipient of the suggestion, only acknowledges the suggestion as °*so ya naa* ° '°yeah°' in a smaller voice. She has not fully accepted the suggestion.

Aiko's minimal acknowledgment occurs at the same time as Yumi's production of the SSW *buwan*, which is followed by *to suru yatsu* 'the one which goes *buwan*' (line 35). Syntactically, this unit parallels the one deployed when initially drawing the objects. This initial drawing dates back to *ii yatsu* 'one which is good' (the translation keeping the original Japanese word order) or 'a good one' (line 16). However, *ii* 'good' carries abstract meanings. That is, upon the production of the unit *ii yatsu*, Yumi was not explicit about how exactly the thing is good.

Yumi specifies the quality of being good by deploying the sound symbolic unit *buwan to suru yatsu* 'the one which goes *buwan*' that accompanies the gestural aspects of her vocalization and hand gestures. In particular, the speaker Yumi deploys the SSW as if she were counting on its depictive power of vocalization. The SSW *buwan* is vocalized with more air pressure, making a slight explosive sound in /b/, than the surrounding talk, thus audibly displaying the pad's bouncy feature. Further, Yumi keeps her right wrist bouncing up and down; the peak of her movement is synchronized with the middle sound *wa* in *buwan* (Figures 4.20 and 4.21). That is, by coordinating the available semiotic resources—language, prosody, and hands—Yumi concretizes the object (the wrist pad/rest) while demonstrating how it works in the material

environment. Figure 4.21 also shows that, at the end of the sound symbolic unit, Yumi holds a mutual gaze with Aiko, indicating that Yumi selects Aiko as the recipient of the object concretization.

After hearing the unit that incorporates the SSW, Aiko accepts the suggestion by verbalizing *un wakatta* 'yeah, I understand' (line 38), followed by her plan *katte morau wa* 'I'll ask him to get one for me' (line 39). In sum, Yumi offered suggestions three times, the third of which she deployed the SSW to depict the property of the object drawn through the deployment of the demonstrative *kore* 'this one.' As a result, Yumi finally received a sequentially preferred response (i.e., acceptance).

This segment has demonstrated that the speaker counts on the possibility of SSWs in depicting the properties of the object. Even after producing the visual display of the object, the speaker provided the recipients with the audible depiction through the deployment of the SSW. In contrast, the next segment shows that SSWs are not always capable of depicting properties of objects. Such a limitation invites the speaker to reformulate what is expressed via SSWs by adding non-sound symbolic verbal expressions.

The following segment comes from a three party talk among friends: Aki, Sanae, and Mio. Prior to the segment below, Aki talked about why she had broken up with her boyfriend. In her response, Sanae displayed her understanding of Aki's circumstances by summarizing what Aki had said, but Aki did not exactly confirm Sanae's summary. Aki further explained that her ex-boyfriend cared a lot about their sexual relationships, something she says she empathizes with. The SSW occurs while Aki explains how her attitude toward their sexual relationships differed from her ex-boyfriend's attitude (lines 19 to 24). The analysis first focuses on how Aki draws the object—a hypothetical scale—into the interaction. Aki then concretizes the object by deploying

the SSW that depicts the property of the object—the force of her ex-boyfriend's sexual desire— , but ends up facing limitations.

(4.6) [WKR.5006]

19 Aki: *tada:* (0.6) *nanteyu:no*
even.so whatchamacallit
'even so (0.6) whatchamacallit'

Aki

Figure 4.22

Figure 4.23

* *

20 Aki: *kore ga sa: shakudo da to shite,*
this.one SB FP scale CP QT do.and
'this is the scale, assuming so, and'

Figure 4.24

Figure 4.25

Figure 4.26

* * *

21 *kare ga kon dake furu de sa,*
he SB this much full with FP
'he is this much, with full'

Figure 4.27

22 → *bua::* *tto* *mo:* [°*sono* *sa*°
SSW QT y'know that FP
'like *bua::*, y'know, like'

23 Sanae: [° *mh* *hhu*°
'mh hhu'

Figure 4.28

Figure 4.29

24 Aki: *furi*[*ki(h)re(h)* *cha(h)u* *gu(h)rai*[*ni* *sa*
off.the.scale done degree FP
'it's off the scale, to such an extent'

25 Mio: [*huh*

26 Sanae: [*hah* *ha*[*h* *hah* *ha*

27 Aki: [*ha* *hah* *hah*

In line 19, Aki restarts her turn with *tada:* 'even so,' which indicates that some condition would have to follow. But what follows is a 0.6-second pause and *nanteyu:no* 'whatchamacallit,' which displays her trouble with wording something. Here, Aki starts to prepare the forthcoming hand gesture: she brings her hands in front of her body, presses them together, and starts gazing at them. Her eyes do not meet with any of the co-participants' eyes until line 28, when Sanae provides a verbal response (not included in the segment).

After indicating trouble, Aki draws the hypothetical scale into the interaction to explain her ex-boyfriend's sexual desire as follows. While producing *kore* 'this one' (line 20), Aki brings her palms together (Figure 4.22). Her verbal production of *ga sa:* in *kore ga sa:* 'this is' is accompanied by a widening movement of her hands to shoulder-width (Figure 4.23). While maintaining the hand-width distance, she brings the hands down on the table, identifying what she gestured with her hands as *shakudo dato shite* 'the scale, assuming so, and.' Here, the hypothetical imaginary scale comes into being through the coordination of various multimodal resources: the demonstrative *kore* 'this one,' the hands designating width, the lexical resource *shakudo* 'scale,' and the grammatical construction *dato shite* 'assuming so, and.' It should be noted that the element at the very end, 'and,' indicates that there is a continuation.

The hypothetical imaginary scale then serves as a situated resource for Aki to continue the ongoing turn. In line 21, Aki brings her right palm to the left while saying *kare ga* 'he' (Figure 4.24). By doing so, Aki indicates that she is going to measure something related to the ex-boyfriend, most likely his attitudes toward sexual relationships, although she has not clearly stated it yet. She again widens and narrows the hand-width gap while producing *kondake furu desa* 'this much, with full' (Figures 4.25 and 4.26). In this unit, the occurrence of another demonstrative expression *kondake* 'this much' is synchronized with the widening movement of the hands, which invites the recipients' gaze.

Hers, Aki's hand width is maximized, at which point the adverb *furu de* 'with full' is produced. This hand movement visually enhances the projection of the type of the forthcoming element. Aki then rapidly narrows the hand-width while producing the clause final elements *de sa*, indicating that the turn has not been completed yet but continues (Figure 4.26). This clause is followed by the sound symbolic unit *bua:: tto* 'like *bua::*' that occurs simultaneously with a

swift widening movement of her hands (Figure 4.27). It is remarkable that the sound <u>*bua::*</u> is produced with a stronger breath than the surrounding talk, while the vocal gestural aspect of the SSW manifests itself. By juxtaposing words, prosody, and hand movements, Aki acts out the property (i.e., the intensity of sexual desire) across the hypothetical scale. The sound symbolic unit is then followed by *mo: sono sa* 'y'know, like,' while Aki gazes at her hands, indicating that there is more to come.

Let us consider the recipients' responses in order to see how they understand Aki's turn that incorporated the SSW. Upon hearing the sound symbolic unit, Sanae only produces a laughter token °*mh hhu*° in a small voice, while Mio displays no particular response other than gazing at Aki. These recipients' minimal responses are most likely due to their uncertainty regarding what the speaker had depicted. The object –the scale—was introduced in the interaction but operated as an analogy. Building on the analogical scale, Aki depicted the property through the SSW, but had not clearly stated what she was going to depict through the SSW up to this point. That is, what the sound depicted is not crystal clear to the recipients upon the vocalization of the SSW.

The lack of a clear display of understanding by the co-participants makes further clarification relevant, in the speaker's pursuit of a preferred response. Subsequently, Aki produces an alternative formulation of her own original turn, which further reveals that deployment of SSWs does not always operate as a means of depiction, but the speaker can also face some limitations. In line 22, by producing *sono sa* 'like,' Aki indicates trouble with continuing the turn. Aki then produces a reformulation that consists only of non-SSWs, *furikirechau gurai ni* 'it's off the scale, to such an extent,' in a laughing voice (line 24) (Figures 4.28 and 4.29). While doing so, she recycles the hand movement she made by widening and

narrowing her hand-width. By means of this hand movement, this reformulated unit is visually designed to tie with the prior sound symbolic unit *bua:: tto* 'like *bua::*.' The reformulated unit contains the verb *furikireru*, which in Japanese exclusively describes the state of the pointer going off the scale. Thus, upon hearing the initial part of the replaced unit, *furikirechau guraini* 'it's off the scale, to such an extent,' Mio is able to display her understanding by bending her upper body forward without speaking audibly. The other recipient Sanae bursts into laughter after hearing the verb. What allows both recipients to show their understanding of Aki's course of the turn is not the SSW, but the reformulated turn that consisted of non-SSWs. A lack of recipients' display of understanding led to the expansion and reformulation of the turn.

This section has disclosed the possibilities and limitations of SSWs in depicting properties of objects. In particular, the objects examined in this section have been ones newly made in interaction by means of the demonstrative pronoun *kore* 'this one' and co-occurring hand gestures. The next section focuses on SSWs in word search activities, where both possibilities and limitations intersect with one another.

4.4 SSWs as a provisional element in word search activities

Owing to the capacity of visual resources and speech to occur simultaneously, deployment of SSWs contributes to concretize objects instantly. This section examines deployment of SSWs after the speaker indicates trouble with continuing the ongoing turn. The speaker indicates this trouble by producing, among others, sound stretches, intra-turn pauses, and word cutoffs. Hayashi (2003b), who examined word searches in Japanese conversation, stated that these delaying elements "may co-occur with, or be followed by, several recognizable features" as in (4.7):

(4.7)

1. "Delaying device" like *ano*: (uhm), *nanka* (like), and so forth
2. Self-addressed questions for recollection like *nan da(tta) kke* (What is/was it), and so forth
3. Orientational shifts, typically aversions of gaze away from the addressee (cf. Goodwin & Goodwin, 1986, for word searches in English)
4. A variety of manual and facial gestures, including iconic gestures that represent some aspect of the searched-for item, as well as a characteristic "thinking face" (cf. Goodwin & Goodwin, 1986, for word searches in English)

(Hayashi, 2003b:113)

After indicating trouble with continuing the turn, the speaker may produce a searched-for item. One of the typical searched-for-items is an object name, which could alternatively be depicted through an SSW that concretizes an object. This section demonstrates how SSWs operate in the process of word-search activities. More precisely, it demonstrates that when the speaker deploys SSWs after indicating having had trouble with continuing the turn, whose design shows the following characteristics: The two individual phases (drawing objects into the interaction and adding properties to them) become integrated into one, due to the co-occurring nature of vocal gestures and hand gestures. By doing so, the speaker pushes the interaction forward.

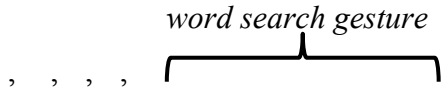
4.4.1 Outlining word search activities

Before examining cases of word searches, one point of possible confusion must be addressed. In the previous studies, there seems to be no clear statement about the size of the grammatical unit of the searched-for item. One of the typical formats of word searching turns is "What was X?," in which X refers to the name of a thing, place, person, or else. With this type of word searching turn, the projected searched-for item adopts the form of a noun. This chapter, however, does not limit the possible occurrence of the searched-for item to just one-word. Over the course of the interaction that involves word searches, the speaker's deployment of various resources projects

the forthcoming occurrence of a "descriptive element of some sort" (Hayashi, 2003b:115). As a consequence of the search, the element that was produced as the searched-for item could be, for example, a noun phrase, or more specifically, a verbal modifier followed by a noun. An example is *ase kakitai hi?* 'The day when ((you)) want to sweat?,' produced as a searched for item (Hayashi, 2003b). Labeling the target activity "word search" does not necessarily mean that the searching is only for one word, but for a unit that may be larger than a word.

Having clarified the type of grammatical unit that can occur at the position in which the searched-for item is due, we can now outline how the verbal and bodily resources listed above (4.7) are deployed, and how they work in face-to-face interaction. The speaker may do a "solitary search" (Goodwin, 1986:63), in which speakers search for a word, find it, and then produce it themselves. Or, the speaker may start a search, and the other participants might collaboratively help them find a word. This section presents two segments of data. In both cases, the speaker seems to initiate a collaborative search, but ends up producing an SSW as a provisional element by themself. The following summary primarily focuses on studies of solitary searches in order for the reader to grasp the basic idea of how word searches occur. The following short segment displays the essence of solitary word search activities.

(4.8)



Speaker : he pu:t uhm, (0.7) tch! put crabmeat on th'bo::dum.

(M. Goodwin, 1983:129-30)

This example consists of several linguistic practices that indicate that the speaker is unable to produce an appropriate word right away, and is thus searching for a word. The speaker first

indicates that he is having trouble with finding a word by prolonging the vowel in the verb *pu:t*, which is followed by the production of *uhm* and a 0.7-second silence, both of which also indicate the speaker's trouble with continuing the turn. That is, the searched-for item is the object of the verb put (i.e., the thing that was put on the bottom). Then, the speaker's "self admonishment" (p. 129) *tch!* follows. Non-vocal display also indicates that the speaker engages in searching for a word. The commas printed above the utterance of *he pu:t uhm* indicate that the speaker is gazing away from the recipient. This way of using the gaze—taking away from the recipient—is a common practice when the speaker engages in solitary search. In contrast, by bringing the gaze to the recipient, the speaker solicits a help from the recipient (Hayashi, 2003b). Going back to Goodwin's example, the speaker then makes "a thinking face" (p. 130) as annotated as a "word search gesture" in the segment. What follows is the reproduction of the verb *put,* which is, this time, not accompanied by a sound stretch. Without pausing, he produces the searched-for item that was *crabmeat*. The following flow chart summarizes the word search sequences discussed in previous studies:

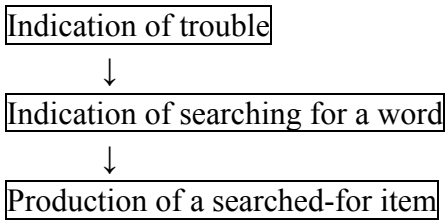


Figure 4.30. Basic Word Search Steps Based on Previous Studies

The deployment of SSWs in word search sequences alters these steps. The target steps examined in this section are schematized in Figure 4.31 below. As Section 4.4.2 will demonstrate, two primary interactional characteristics stand out. The distinct stages are printed in italics:

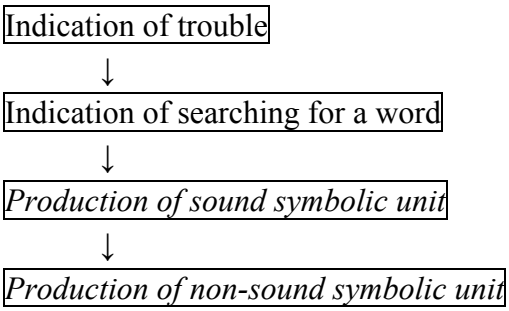


Figure 4.31. Word Search Steps in Which SSWs Are Deployed

One of the characteristics involved in the deployment of SSWs in word search sequences is seen in turn designs. Upon the vocalization of the SSW, the speaker draws the object into the interaction while he/she depicts the property of this object. By simultaneously performing the two phases—drawing the object and adding its property—the speaker is able to push the talk forward. The other characteristic is that the sound symbolic unit (i.e., an SSW followed by other grammatical elements such as a quotative particle and a verb) will then be replaced with another unit, which exclusively consists of non-sound symbolic elements. For this reason, this chapter claims that SSWs serve as provisional elements in word search activities. The next section examines two cases of word search activities that exemplify both characteristics.

4.4.2 How SSWs operate as a provisional element in word search activities

The following segment comes from a talk among four close friends at a university in Japan. The participants are Risa, Shota, Momo, and Nao. They have been talking about their common friend, who had attended a summer intensive course in preparation for the driver's license exam the previous year. In the segment, the participants (mainly Momo and Shota) search the location of the school, which only Momo knows. In line 11, Nao's WH-question, directed to Momo, requests information about the location. This question opens up a sequence in which the

participants engage in the activity of identifying the location of the driving school. The SSW *tikin* appears in line 12, while Momo engages in describing the location.

Before moving onto the analysis, it is useful to clarify the approximate location of the school, which might help the reader follow the upcoming analysis. As Momo mentions in line 3 in the following segment, the driving school is somewhere in the Fukushima prefecture. Figure 4.32 shows a map of Japan in which the Fukushima prefecture is pointed with an arrow.

Figure 4.32. Map of Japan

To identify the location in Fukushima, the participants refer to the neighbor prefectures of Ibaraki and Tochigi (see Figure 4.33). The arrow in Figure 4.33 indicates the location that Momo describes when deploying the SSW *tikin.*

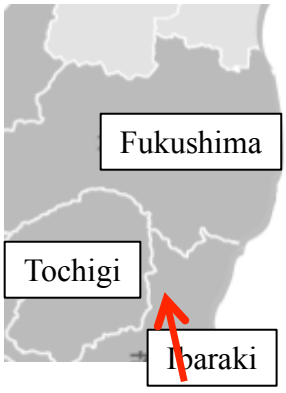


Figure 4.33. Enlarged Map of Fukushima Area

The following analysis first focuses on how Momo's turn unfolds, while she indicates trouble with continuing the turn and searching for a word that would specify this location. This process contextualizes the forthcoming production of the SSW. Then the analysis demonstrates how the object is drawn into the interaction while the property is depicted through the deployment of SSW. The analysis also focuses on how the recipients Shota and Risa attend to the turn that incorporates the SSW.

(4.9) [SKR.1109]

1 Nao: [(e) *doko itta no?*
where went FP
'where did she go?'

2 (0.6)

3 Momo: *Fukushima.*
'Fukushima'

Figure 4.34

Figure 4.35

* *

4 (1.0)

Figure 4.36

*

5 *tte tteta. (.) ano::, ichiban, tochigi yori no*
QT saying.PT uhm the.most Tochigi close LK
'she said. (.) uh::m, the closest to Tochigi'

6 (1.0)

7 Shota: ((small nodding))

Figure 4.37

*

8 (1.2) ((Momo and Shota sustain mutual gaze))

Figure 4.38

*

Figure 4.39

*

9 Shota: *˚ichiban˚ tochigi yori? shita no [hoo ˚da yo ne˚*
the.most Tochigi close down LK side CP FP FP
'the closest to Tochigi? It's down there, isn't it.'

10 Momo: [*shita no hoo no*,
down LK side LK
'down there'

11 (.)

Figure 4.40 *Figure 4.41* *Figure 4.42* *Figure 4.43*

* * * *

12 → *ano:, Ibaraki no-* (.) *tikin te natteru toko no*
uhm Ibaraki LK SSW QT becoming place LK

'uh:m, in Ibaraki (.) the place where it's being *tikin*'

Figure 4.44

13 (.) *kocchi gawa*( )
this side
'this side of it'

14 Shota: [((*nodding*))

15 Risa: [°*ti(h)kin te na(h)tteru t[te*°
SSW QT becoming QT
'you said bing *tikin*'

16 Momo: [*iba[raki ga kuikonde[ru* toko=
Ibaraki SB digging.in place
'the place where Ibaraki is digging in'

17 Nao: [°(*tiki(n)ka*)°
SSW Q
'*tikin*'

18 Risa: [hh

19 Momo: =*aru (h) ja nai desu ka.*
exist CP NG CP Q
'there is such a place, isn't there'

20 Shota: ((*nodding*))

21 Momo: *no[::, kocchi gawa.*
LK this side
'this side of it'

As already mentioned, Nao in line 1 requests information about the location. However, a 0.6-second pause takes place, which delays the production of the answer (line 2). Momo then provides an answer with the prefecture name, which is rather a big category with which to refer

to a very specific location (line 3). Even after the answer is provided, a 1.0-second silence takes place (line 4). During this silence, Nao, who initiated the question-answer sequence, could have produced a sequence closing third; however, he first produces a small nod and then tilts his head (Figures 4.34 and 4.35), displaying that he does not fully understand the location based on Momo's answer.

The display of his lack of understanding leads Momo to continue her turn to further clarify the location as follows. Momo expands her turn by attaching *tte teta* 'she said' (line 5). Once again, a micro pause takes place, and Momo continues her turn with a delaying item *ano::* 'uhm' (line 5), which is stretched, indicating that Momo is searching for a word. In the meantime, Momo prepares for the forthcoming hand movement: she brings her right index finger to the top of the table, draws something like a circle, and then traces it with her finger repeatedly (Figure 4.36). Momo further continues her turn with *ichiban tohcigi yori no* 'the closest to Tochigi' (line 5) and narrows down the area in Fukushima. We should note that the phrase 'the closest to Tochigi' relates to the distance between Fukushima and Tochigi.

Momo's turn still does not receive any responses from any participants. Noticeably, a 1.0-second silence takes place (line 6). Shota then produces small nodding (line 7), but other recipients show no response at all. A 1.2-second silence follows then, which is relatively long. During this silence (line 8) it is remarkable that Momo is still tracing the imaginary circle on the table, while Momo and Shota sustain a mutual gaze. By doing so, Momo solicits Shota's help in finding a word (Figure 4.37).

At this point, Momo and Shota start to draw the object into the interaction by co-constructing the imaginary map in the following way. Shota repeats Momo's phrase *ichiban tochigi yori* 'the closest to Tochigi,' which ends in a rising intonation (line 9). While repeating

the phrase *ichiban tochigi yori* 'the closest to Tochigi,' Shota looks up in the air (Figure 4.38). Then he provides an alternative way of viewing locations on the map, *shita no hoo da yo ne* 'it's down there, isn't it,' which accompanies the following hand gesture: he places his right palm facing down in front of his body, and moves it up and down (Figure 4.39). Along with this hand gesture, he produces head nods which claim his understanding of the approximate location that Momo has described. Shota's turn proves crucial in that he has changed his way of referring to the locations on the co-constructing imagery map: from indicating the closeness of the two prefectures to remarking on how the two prefectures are adjacent to one another.

In line 10, this time, Momo incorporates Shota's phrase *shita no hoo* 'down there' into her turn by repeating it. This incorporation is followed by the linking particle *no*, which projects another nominal phrase to come (not reflected in the English translation). However, a micro pause takes place (line 11), and Momo produces *ano* 'uhm,' which delays the progression of the turn (line 12). She then continues by introducing another adjacent prefecture *Ibraki no* 'in Ibaraki,' but cuts it off. Then, another micro pause takes place. Up to this point, the way Shota and Momo have constructed their turns indicates that the upcoming element of the turn would be associated with the searched-for location, pictured in their co-constructed imagery map. Further, the location could be identified using the newly introduced prefecture, Ibaraki, as the reference point. However, they do not specify the location's whereabouts in that way.

Instead, building on the co-constructed imagery map, Momo depicts its property by deploying the SSW *tikin*, which is followed by the elements *te natteru toko*, altogether forming a sound symbolic unit: 'the place where it's being *tikin*' (line 12). The production of the sound *tikin* is synchronized with this movement: she brings her finger up quickly to the level of her head and brings it down, as if the finger were sticking to something in the air (Figures 4.40 to

4.43). It is also noticeable that *tikin* is vocalized a little strongly, in a slightly high pitch that makes the quality of the vocal gesture quite noticeable. Momo juxtaposes various semiotic resources to depict the location upon the deployment of the SSW *tikin*, thereby moving the interaction forward.

Although the SSW *tikin* is supposed to depict the location of Fukushima on the hypothetical map, Momo has not successfully done so in a way that elicits a display of understanding from the recipients. Here, Momo depicts the property (i.e., being *tikin*) at the same time she pictures the map (i.e., the exact location of the place being *tikin*). That is, the map was not fully prepared prior to the depiction of its property through the deployment of the SSW. Momo counted on the possibility of the SSW; however, it turned out that the way she designed her turn does not publicly display her understanding of the location. Nevertheless, in line 13, Momo further continues her turn by using the place depicted through the phrase *tikin te natteru toko no* 'where it is being *tikin*,' as a reference point: *kocchi gawa* 'this side of it' (Figure 4.44). When Momo is about to complete the turn, she shifts her gaze back to Shota.

The recipients, Shota and Risa, react to the turn that incorporates *tikin*, but they provide different types of responses. Receiving Momo's gaze, Shota produces a nod, which claims his understanding of the location. On the contrary, Risa, who has been observing the exchange between Momo and Shota, mocks Momo by taking a turn as *tikin te natteru tte* 'you said being *tikin,*' in line 15, in a laughing voice.

Risa's mocking response turns out to solicit Momo's reformulation: *ibaraki ga kuikonderu toko* 'the place where Ibaraki is digging in.' This is followed by a confirmation request as to whether or not her reformulated turn is understandable. In the meantime, Risa shows a smile on her face, and produces a laughter token (line 18). On the other hand, Shota

claims his understanding by nodding (line 20). It is important to note that the two units, before and after the reformulation, share the same syntactic format: a modifier in *teiru* '*–ing*' form followed by the noun *toko* 'place' (i.e., *tikin te natteru toko* 'the place where it's being *tikin*' and *ibaraki ga kuikonderu toko* 'the place where Ibaraki is digging in'). That is, by reformulating the turn in this way, Momo displays her understanding that the turn that incorporated the SSW had a limitation. However, Momo surely pushes the interaction forward by the deployment of the SSW.

The next segment also shows that the speaker counts on the possibilities of SSWs, but turns out facing limitations. The participants of the following segment are Nana and Akira, who are friends. In the segment, Nana proffers the following topic: she received a message from their common friend Misa. Akira then informs her that he also saw the message, and he begins to search for a word. It should be noted that, from the beginning of this segment up to the middle of line 7, where Akira moves his arm, Nana engages in two different activities: interacting with Akira and connecting her smartphone to an adaptor that will be placed on the table afterward. Due to her engagement in two activities, Nana tends to redirect her gaze down toward the smartphone. As it will be shown, Nana's downward gaze might be relevant to Akira's deployment of hand movements, which draws Nana's gaze. Paying special attention to Akira's deployment of multimodal resources, the analysis focuses on how he gradually shifts the configuration of the turn (i.e., an arrangement of bodily and visual elements in his turn), while searching for a word. In this process, Akira produces an object (i.e., smartphone) by acting out one of its properties (i.e., how to use it). As it will be demonstrated, Akira first produces his outcome of the search in line 7, exclusively with hand movements, and then reformulates it into one co-occurring with the SSW *shushushu* in line 9. My analysis also focuses on how the recipient Nana monitors the configuration shift of Akira's turn and reacts accordingly.

(4.10) [CHH.0027]

2 Nana: *iya sa, kesa renraku kita wake yo,*
well FP this.morning message came reason FP
'well, this morning a message came to me'

3 *a[no::*
uhm
'uhm'

4 Akira: *[°nani ga?°=*
what SB
'°what's that? °'

5 Nana: *=misa kara*
((personal name)) from
'from Misa'

6 (0.8)


Figure 4.45

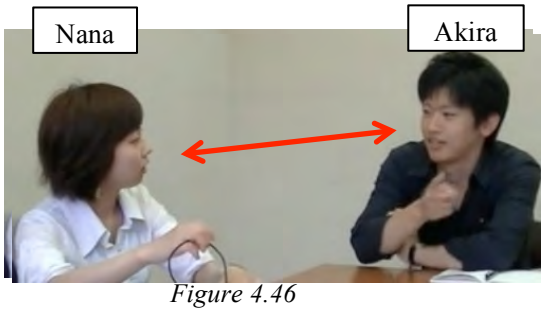


Figure 4.46

* *

7 Akira: *ima mita ano:* (0.8)
now saw uhm
'I just saw that now uh:m'

8 Nana: *ah: [rain?*
oh Line ((smartphone app name))
'oh *Line*?'

Figure 4.47

*

9 Akira: → [*shushushu tte yatte*
SSW QT do.and
'I did like *shushushu*, and' ((keeping the Japanese word order))
'by doing like *shushushu*' ((more natural word order in English))

10 (0.6)

11 *sokuhoo*, *sokuhoo¿*=
newsflash newsflash
'newsflash, newsflash?'

12 Nana: =*ah(h)a* [*so(h)ku(h)ho:hh*
oh newsflash
'o(h)h, ne(h)wsfla(h)sh hh'

13 Akira: [*nanka, deru jan.*
something appear TAG
'something, appears huh?'

14 Nana: *hh sokuhoo tte*
newsflash QT
'hh you said newsflash'

15 Akira: *ya nanka*
no like
'well, like'

16 Nana: *n, ya, kor-, kore desho?*
umm no thi- this.one TAG
'umm, no, thi-, this one right?'

17 ((Akira and Nana look into the screen of Nana's smartphone))

18 Akira: *soo.*
so
'right.'

Responding to Nana's topic proffering (line 2), Akira requests a clarification (line 4). Nana then informs Akira that the message was from Misa. Having discovered that the message was from Misa, Akira responds as: *ima mita ano* 'I just saw that now uh:m' (line 7), ending with the element *ano* 'uhm.' This ending indicates that he is going to continue the turn, but has trouble with doing so. This is followed by a 0.8-second silence, at the beginning of which he starts moving his right arm and, with it, draws Nana's gaze.

Having focalized Nana's gaze, Akira bodily displays that he is searching for a word: He brings the right index finger in front of his body. He then makes quick and small repetitive movements from side to side while resting his left arm on the table (Figure 4.45). By virtue of the sequential position, this hand movement seems to create an object related to the fact of having seen the message. Akira, at the same time, indicates that the upcoming element of his turn is something related to those hand movements. While Akira produces the hand movements, Akira and Nana hold gazes (Figure 4.46). Akira's use of the gaze indicates that he is soliciting help from Nana in finding a searched for word (Hayashi, 2003b).

The salient feature of Akira's turn configuration is that, during the 0.8-second silence, his bodily movement occurs, but no verbal conduct takes place (in line 7). That is, Akira started producing the turn but he himself interrupts it via his hand movements, thereby indicating that he is already in the process of searching for a word. Akira's bodily conduct is visible to Nana, who then shows her understanding that Akira has been engaging in searching for a word associated with his having seen the message. In line 8, Nana displays her noticing by producing *ah:* 'oh' (Heritage, 1984), based on her observation of Akira's turn. She then provides an aid for Akira by offering a candidate item, *rain* 'Line' ending with a rising intonation.

However, before Nana offers the candidate item, Akira comes in and produces an outcome of the search by deploying the sound symbolic unit: *shushushu tte yatte* 'I did like *shushushu*' (line 9), with accompanying hand gestures (Figure 4.47). Examined more closely, Akira's turn in line 9 can be analyzed as a reformulation of his prior turn during which he produced a similar hand gesture using only his right hand in line 7. While producing this reformulated turn (i.e., *shushushu tte yatte* 'I did like *shushushu*') in line 9, Akira visually restructures the stage for the hand gesture as follows: he straightens up his back and turns his left palm to face towards himself, as if he was holding something in his hand. Having established this left-hand shape, he recycles the prior right hand movements that went side-to-side repeatedly (Figure 4.47). Here, by virtue of the sequential position, the left hand is supposed to model a smartphone, with the right hand mimicking the hand operating it. This revised version of the hand gesture is rhythmically coupled with the verbal production of the SSW *shushushu*, which is followed by *tte yatte* 'I did like.' Additionally, still, Akira's gaze is directed to Nana. While deploying the SSW *shushushu*, he coordinates language, prosody, and gesture, thereby concretizing the object instantly. Differently put, he counts on the possibility of the SSW when drawing the smartphone into the interaction, during which he added a property relevant to this object (i.e., how to use). By doing so, he moves the interaction forward.

However, the SSW appears to demonstrate its limitation. That is, Nana does not immediately appreciate it, as indicated by the occurrence of the 0.6-second silence (line 10). During the silence, Akira and Nana hold a mutual gaze. Following Nana's non-response, Akira replaces the prior production of the sound symbolic unit with the noun *sokuhoo* 'newsflash' (line 11). However, he immediately questions his own selection of the word by repeating it with a

rising intonation as *sokuhoo¿* 'newsflash?.' Such a self-doubtful verbal production reveals that he has been facing trouble in finding a word to fit his turn.

Akira's self-doubtful conduct draws a response from Nana, who displays her noticing by producing another change of state token: *ah* 'oh' in a laughing voice (line 12). She then mocks Akira's word selection by repeating it as *sokuhoo* 'newsflash,' still in a laughing voice (line 12). Akira immediately reacts and overlaps with this turn by further reformulating what he has specified as *sokuhoo* 'newsflash' with *nanka deru* 'something appears' (line 13). By doing so, Akira displays his orientation to the problematic part of his own prior turn, which was *sokuhoo* 'newsflash.' Nana produces a laugh token and again repeats *sokuhoo* 'newsflash' (line 14), followed by the quotative particle *tte*, thereby mocking him. On the other hand, by producing delaying items *ya nanka* 'well like' (line 15), Akira keeps indicating trouble with finding a word that would display something related to how he saw the message from their common friend. Despite the seeming lack of an appropriate word selection, Nana shows Akira the screen of her smartphone (line 16), seeking his confirmation that what he has been depicting is what she shows. Both look into the screen of Nana's smartphone (line 17), and Akira confirms it (line 18). At this point, Akira and Nana reach a mutual understanding.

This section has demonstrated how SSWs operate as a provisional element in word search activities. Notably, the two phases, drawing an object into the interaction and adding the property to the object, have been integrated into one. When the speaker has trouble with continuing the ongoing turn, concretizing an object through the deployment of the SSW operates as a way to push the interaction forward.

4.5 Summary of chapter 4

The analysis of this chapter revealed how absent objects were concretized through the deployment of SSWs in the interactional space. First, it demonstrated that an object was drawn into the interaction by means of demonstrative expressions such as *sore* 'that one,' which referred to the object which had been already talked about. The object then gained properties through the deployment of SSWs. Next, this chapter illustrated how the possibilities and limitations arose while the speaker depicted the property of objects through SSWs. When objects were drawn into the interaction using the demonstrative *kore* 'this one' and co-occurring iconic hand gestures, limitation and possibilities came into play. The speaker may reformulate the prior production of the SSW into non-sound symbolic verbal expressions, especially when the speaker did not receive a response due to the unclear referent. Or, the speaker may produce the SSW even after other semiotic resources (e.g., words and hand gestures) have established the property of the object. Finally, this chapter examined cases in which SSWs were deployed in word search activities. It revealed that the two phases, drawing an object into the interaction and depicting its property via SSWs, were integrated into one when SSWs were deployed as a provisional element. By concretizing absent objects in interaction when having trouble with continuing the turn, the speaker pushes the interaction forward.

Chapter 5

Rendering Emotional Stances into Sound Symbolic Words

5.1 Introduction

This chapter examines sound symbolic words (SSWs) that render the speaker's emotional stance in face-to-face interaction. Emotional stances are not always easy to display with specific adjectives and adverbs. In such cases, SSWs operate as a readily available resource which can render one's emotional stances by nature of their depictive efficacy. As Chapters 2 and 3 demonstrated, the deployment of SSWs is interlaced with multimodal displays of vocalization and the accompanying gestures, and are thus capable of depicting subtle and nuanced emotional stances. In previous studies, emotional stances in interaction were often classified into categories such as anger, indignation (Couper-Kuhlen, 2012; Selting, 2010), annoyance (Couper-Kuhlen, 2012), disgust (M. Goodwin, Cekaite & C. Goodwin, 2012), distress (Wootton, 2012), and surprise (Wilkinson & Kitzinger, 2006). However, my data suggest that immediate emotional reaction to a given situation cannot always be easily classified into a category. This chapter demonstrates that SSWs work as a readily available resource for the speaker to render his/her emotional stance which is not easy to depict with non-sound symbolic words (non-SSWs). It also demonstrates that SSWs work as a resource for the recipient to display his/her understanding of the prior speaker's emotional stance, which has been depicted through an SSW. The analyses of these two patterns further reveal that there is a distribution tendency regarding the sequential positions where conventional SSWs occur and where invented SSWs occur, respectively.

The target class of SSWs in this chapter has been called *gijyoogo* 'psychomimes' which "symbolize[s] mental conditions or sensations" (Shibatani, 1990:154). Some *gijyoogo*

'psychomimes' are conventionalized and listed in dictionaries. They are often followed by the verb *suru* 'do' (Tamori & Schourup, 1999). Examples include *waku waku suru*, "the state of being restless in anticipation of something soon to occur" (Kakehi, Tamori, & Schourup, 1996b: 1252); *doki doki suru*, "the state of being excited or startled by something" (Kakehi, Tamori & Schourup, 1996a:265); and *shonbori suru*, "the state of being sad, lonely or depressed" (ibid:1114). In conversation, as we observed in the preceding chapters, the speaker may utilize conventional SSWs, or newly invented SSWs, whose referents are not necessarily something visible. In the same way, speakers may invent SSWs when rendering their emotional stances.

Studies of emotions from an interactional perspective (e.g., Couper-Kuhlen, 2012; Drew, 1998; Mandelbaum, 1991/1992; Selting, 2010; Sorjonen & Peräkylä, 2012) have established that participants coordinate multiple semiotic resources while constructing turns in which the speaker displays his/her emotional stance. These studies emphasize that verbal, prosodic, and bodily resources are all interlaced and operate as a unified whole while the turn unfolds. Among various multimodal semiotic resources, facial gestures are recognized as visual resources that shape the ongoing interaction (Peräkylä & Ruusuvuori, 2006, 2012; Ruusuvuori & Peräkylä, 2009). In a conversation analytical approach, researchers such as Peräkylä and Ruusuvuori analyzed data from video-recorded conversation and demonstrated that facial gestures ("facial expressions" in their terminology) mark participants' stance toward stories. For example, Ruusuvuori and Peräkylä (2009) focused on facial gestures involved with assessing stories. They demonstrated that a listener's smile indicates his/her understanding of the preceding story as humorous one. Using a different segment, they also demonstrated that a teller's frowning face prior to a telling indicates that the subsequent story will be conveyed as something problematic. Through examinations of laughter and smiles, Haakana (2010) demonstrated that smiling "can be seen as

a pre-laughing device" (p. 1510). Their studies have established that facial displays project the subsequent trajectory of the turn.

In discussing how actions are recognized, Levinson (2012) argued that a turn may simultaneously perform multiple actions among which a "main job" (p. 107) may receive a response from a co-participant subsequently. Displays of speakers' emotional stances could be a job that is "intended to be recognized alongside the primary actions" (Sorjonen & Peräkylä, 2012:9). However, in performing actions such as complaining, assessing, or teasing, "action and emotion can be considered as more closely intertwined" (ibid). This chapter consists of data segments in which the speaker describes his/her firsthand experience,[1] all of which are negative, and include troublesome situations, annoying colleagues, disgusting food, and so forth. Thus, most of the data segments can be categorized as troubles-telling (Jefferson, 1988, 2015). Speaker's subsequent emotional displays through SSWs are therefore associated with frustration, anger, disgust and so forth.[2]

There is a similar line of research that examines complaint stories (e.g., Drew, 1998; Drew & Holt, 1988; Günther, 1997; Haakana, 2007; Mandelbaum, 1991/1992). Scholars found that when constructing complaint stories, the teller provides assessment and displays his/her emotional stance in addition to "reporting talk," which is "the reproduction of prior talk in a current interaction" (Clift & Holt, 2007:1). Haakana (2007) emphasized the importance of examining "reporting thought," or reproduction of prior thought, as a distinct phenomenon from

[1] As summarized in Chapter 2, Kita (1997) claimed that SSWs convey "affecto-imagistic dimensions" of meaning, which represent "the affective, emotive, and perceptual" (p. 387) aspects of experience. Kita's (1997) finding on meanings of SSWs seems to coincide with a sequential pattern observed in this section. In the conversation segments examined in this chapter, SSWs are produced when the speaker depicts his/her emotional stance which originates from his/her firsthand experience.

[2] In my database, the participants tend to talk about something negative. If participants talk about something positive, for example, happy experience, they might have produced SSWs that render happy cheerful emotions.

reporting talk. He demonstrated reporting thought as a component of complaint stories. In my data, what SSWs accomplish resembles reporting thought in that the speaker informs the recipients of their own past thoughts, or more precisely emotional stances. Reporting thought differs in that SSWs are vocalized in a way that the teller acts out the emotional stance in the current interaction. As previous chapters, especially Chapter 3 on SSWs as vocal gestures, have demonstrated, speakers manipulate the phonetic shape of the SSW, thereby vocally gesticulating. Furthermore, speakers also produce body movements during the vocalization. By juxtaposing various multimodal resources, the speaker displays or even enacts his/her emotional stances in the current interactional space. Through enacting themselves or even others, speakers represent what they have experienced somewhere sometime in the past as if it were happening in the ongoing here-and-now interaction (Sidnell, 2006).

This chapter is organized as follows. Section 5.2 examines deployment of SSWs after the speaker describes scenes from his/her negative firsthand experience. It shows that the speaker deploys SSWs to render his/her emotional stance, which is not easy to depict through non-SSWs. In face-to-face interaction, recipients can also experience the speaker's emotional display and orient to it. Section 5.3 demonstrates that the recipient displays his/her understanding of the prior speaker's emotional stance, through production of the same or a phonetically similar SSW as the speaker does or has done. It appears that the recipient had knowledge of the speaker's/teller's troublesome situation to some extent. These analyses of SSWs occurring in different sequential environments suggest that new, or invented, SSWs tend to occur when speakers depict emotion that originates from his/her firsthand experience, while conventional or shared SSWs tend to occur when recipients show their understanding of the prior turn. Section 5.4 concludes this chapter by summarizing its main findings.

5.2 SSWs as a readily available resource for rendering emotional stances

This section demonstrates that SSWs operate as an interactional resource for rendering emotional stances, which are not easy to depict with non-SSWs. It shows that the speaker faces trouble with finding a linguistic resource that fits depicting their emotional stances, but handle this trouble by vocalizing an SSW. This section presents analyses of three segments. In the first segment, the recipients indicate that the emotional stance depicted through non-sound symbolic adjectives causes a problem of understanding. The speaker then accounts for the scene, and produces the SSW that depicts her emotional stance to assess the referent. By doing so, the speaker indicates that the emotional stance in question was not easily describable with non-SSWs. The recipient in the second segment also indicates a problem of understanding after the speaker rendered her emotional stance. However, the emotional stance in this case is first rendered through SSWs. The speaker then reformulates the sound symbolic unit with a non-sound symbolic unit, thereby leading the participants to achieve intersubjectivity, or mutual understanding of each other's action. The third segment shows a related phenomenon, in which the speaker deploys various SSWs in depicting emotional stances. Due to the instant depiction that SSWs make possible, the speaker realizes emotional transitions from one to the other in a swift manner. The analysis incorporates the concept of "emotional scaling" (Prior, 2016), which will be explained before presenting the relevant data analysis.

The first case comes from a talk among Kaori, Yui, and Nao. Kaori and Yui are college students. Nao is a graduate student, but attended the same program at the same college a couple of years before the data-recording. Prior to the segment, the participants have agreed that the Phonetics course at that college was difficult. In this segment, Kaori tries to deliver another assessment about the Phonetics course, initiating her turn with *ato* 'and' in line 2. The segment is

divided in two parts. The first part (5.1) demonstrates that Kaori depicts her emotional stance using two different adjectives that, in turn, lead Yui to produce repair initiations. Kaori's unsuccessful deployments of two different adjectives indicate that the target emotional stance is neither describable with one word nor categorizable as one specific type of stance. The second part (5.2) demonstrates how Kaori reformulates the turn, by first setting the scene and second deploying the SSW that renders her emotional stance, thereby assessing the referent.

(5.1) [SHR.0429]

1 (0.6)

2 Kaori: *ato nanka* (0.8) *nanka* (1.0) °*nanka* *kimochiwarui*°
and somehow somehow somehow gross
'and it's like (0.8) like (1.0) °like gross°'

3 (0.4)

4 *n(h)n*[*ka h,*
somehow
'like'

5 Yui: [*n? n?*
'what? what?'

6 Kaori: *na*[*nka,*
somehow
'like'

7 Nao: [*ki(h)mo(h)chi(h)*[*wa(h)ru(h)i .h* [*.h*
'gross'

8 Yui: [*eh:(h)?*
'eh:(h)?'

9 Kaori: [*kowai* [*nanka=*
scary somehow
'it's scary, like'

10 Yui: [*eh:(h)?*
'eh?'

In this short segment, Kaori assesses the Phonetics course as *kimochiwarui* 'gross' (line 2). However, immediately after the assessment a 0.4-second silence takes place, which delays the recipients' response. Kaori continues her turn with *nanka* 'like' but one of the recipients, Yui, overlaps with it by seeking clarification, *n? n?* 'what? what?' (line 5). Kaori then replaces the adjective *kimochiwarui* 'gross' with another adjective *kowai* 'scary' (line 9); even after this replacement, Yui still shows that she has not understood what Kaori is trying to describe (line 10). It appears that Kaori is having trouble with finding a linguistic resource that fits depicting her emotional stance toward the Phonetics course in a way the recipients understand.

The continuation of this segment is shown below. The analysis focuses on how Kaori's turn unfolds, during which she describes the scene she experienced in the past. Kaori then provides an assessment of the experience by successfully rendering her emotional stance through the deployment of an SSW. The SSW in this segment is the conventional SSW *zawa*, defined as *samuke nado ga isshun okoru sama* 'the manner in which a chill or shiver occurs for a moment' (Ono, 2007:158).

(5.2) [SHR.0429-2]

11 Kaori: *=minna shite nanka,*
everyone SB somehow
'everyone does like'

12 *sugoi sa yuu jan sensee ga sa yutte sa,*
very FP say TAG teacher SB FP say.and FP
'say something all together, huh? The professor says something and'

13 *minna mo nanka mane suru jan*
everyone too somehow imitate do TAG
'everyone repeats what the professor said, huh?'

Figure 5.1

Figure 5.2

14 *nanka koko .hh [o: mitaina (.) nan[ka are,*
somehow here 'o:' like somehow that
"like here .hh like 'o:,' or like that"

15 Nao: *[ah: : : : : : : [: :*
'I see:::::::'

16 Yui: *[°hai°*
yes
'yes'

Figure 5.3

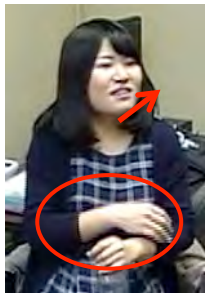


Figure 5.4

Figure 5.5

17 Kaori: → *are ga,.h zawa tte shi[te s- na(h)n(h)ka .h*
that SB SSW QT do.and somehow
'that .h makes me feel *zawa*, and like .h'

18 Yui: *[eh::↑:::?*
'really?'

Starting in line 11, Kaori describes a particular scene that she experienced in the course. First, she introduces *minna* 'everyone,' which refers to the classmates, into the interaction. She then presents the physical activity 'everyone' engages in, as a response to what the *sensee* 'professor' does: *sugoi sa yuu jan sensee ga sa yutte sa* 'say something all together, huh? The professor says something and' (line 12) and *minna mo nanka mane suru jan* 'everyone repeats what the professor said, huh?' (line 13). Without receiving any verbal or physical responses from the recipients, Kaori starts introducing a specific example using her hand and face (line 14). While saying *nanka koko* 'like here,' Kaori brings her left hand close to her mouth. While inhaling, Kaori's mouth makes an o-shape; she touches her cheeks with her fingers (Figure 5.1). This deployment of hand and face continues until she produces the sound 'o:.' By marking the preceding production of 'o:' with *mitaina* 'like,' Kaori indicates that the preceding demonstration was an example of the classroom activity in the Phonetic course. Meanwhile, Nao displays that he understands what Kaori has been describing (line 15). Yui just produces a continuer *hai* 'yes' (line 16).

Kaori then refers to her prior demonstration as *are* 'that one.' The demonstrative pronoun *are* 'that one' is used when the referent is shared among the participants (Kuno, 1973). That is, Kaori's deployment of *are* 'that one' displays her understanding of the referent having been established up to this point. In line 17, Kaori reproduces *are* 'that one.' This time, Kaori marks *are* 'that one' with the subject marker *ga*, thereby indicating that she is going to lead up to some features of the event. She then audibly inhales. At this moment, Kaori raises her shoulders upward, while starting to move her arms (Figure 5.3). Right after this, Kaori crosses her arms, while bringing her chin slightly up (Figure 5.4). These body movements are synchronized with the vocalization of the SSW *zawa*. The SSW is followed by the quotative particle *tte* and the

continuing form of the verb *shite* ‘do and,’ indicating that the preceding SSW displayed Kaori’s emotional stance. As soon as Yui hears the first portion of the verb, she produces a surprise as *eh:↑:::?* ‘really?.’ That is, Yui finally displays understandings of what Kaori has meant by *kimochiwarui* ‘gross’ and *kowai* ‘scary’ in assessing the Phonetics course after hearing the sound symbolic unit.

This segment has shown that the SSW operates as a resource for assessing the past experience that the speaker has described, thereby inviting a display of understanding from the recipient. Kaori’s first try (i.e., *kimochiwarui* ‘gross’) and second try (i.e., *kowai* ‘scary’) were not able to invite Yui’s understanding of what the Phonetics course was like for Kaori. But the third try by means of the SSW successfully received the response from the recipient. From Kaori’s point of view, her emotional stance about the course was not straightforward—something that cannot be fully described through using two adjectives. In contrast, recipients heard two adjectives and then Kaori started describing the scene. That is, the two adjectives worked as the introduction of the telling. The third try of the emotional display through the SSW led the recipient to display her understanding.

The speaker in the next segment also describes a past event, and deploys just such a readily available sound symbolic resource to depict her emotional stance in a given situation, which she has described. However, the depicted stance receives a repair initiation by the recipient. The SSW in question is *gyo*, defined as “the state of being shocked or startled” (Kakehi, Tamori & Schourup, 1996a:496). Other dictionaries register a very similar definition; its meaning has been conventionalized. Having asked for clarification, the speaker reformulates the sound symbolic unit with a non-sound symbolic unit. The design of the reformulated turn

indicates that the teller's immediate emotional stance cannot easily be depicted through non-SSWs.

The participants in the following segment are Yuko and Kimi, college students in Japan who are temporarily staying in the United States. They have been talking about what kinds of toppings are used for pizza, comparing those used in Japan with those in the United States. The SSW appears as Kimi shows her reaction to food that has olives in it (line 7). The analysis focuses on the process through which Kimi indicates trouble with finding a word, concurrently incorporating head movements and facial gestures (e.g., Peräkylä & Ruusuvuori, 2006). This bodily conduct, along with the turn development up to this point, displays that the upcoming component of the turn is something related to her emotional stance in the given situation. The analysis also focuses on Kimi's design of the reformulated turn, demonstrating that the emotional stance is not straightforward to depict.

(5.3) [OKS.1102]

1 Kimi: *watashi* (.) *oriibu toka*,
I olive etc.
'I (.) olives,'

2 Yuko: *a:* [*a:* *a:* *a:*
yeah yeah yeah yeah
'yeah yeah yeah yeah'

3 Kimi: [*minna* *kocchi ireru janai desu ka*
everyone here put TAG CP Q
'everyone here uses olives, right?'

4 Yuko: ((small nods))

Figure 5.6 *Figure 5.7*

* *

5 Kimi: ˚*nanka*˚ (0.6)
like
'like'

Figure 5.8

*

6 → *gyo tte kanji*=
SSW QT feeling
'*gyo* is how I feel' ((keeping the Japanese word order))
'I feel like *gyo*' ((more natural word order in English))

7 Yuko: → *=gyo(h)tte kanji(h)¿*
SSW QT feeling
'do you feel like *gyo*?'

8 Kimi: *sonnani: (.) ireru ka: mitaina*.
that.many put Q like
'like, do you put that many olives on things?'

9 Yuko: *↑a: sokka:, oriibu.*
oh I.see olive
'oh I see, olives'

In lines 1 and 3, Kimi seeks confirmation of her observation *minna kocchi ireru janai desu ka* ‘everyone here uses olives, right?.’ By means of this turn design, Kimi suggests that this observation has already been shared between Yuko and herself. In fact, Yuko comes in right after Kimi started this turn, thereby claiming her understanding of Kimi’s turn (line 2). Simultaneously, Yuko produces a series of small nods, holding a mutual gaze with Kimi. In doing so, she lets Kimi continue her turn. Up to this point, the interaction moves on without delay.

However, at this moment Kimi gazes away from Yuko. This bodily conduct indicates that Kimi is having trouble with continuing the turn. From this point onward, Kimi’s turn grows stagnant. In line 5, Kimi delays her turn by producing ˚*nanka*˚ ‘like’ in a small voice. A 0.6-second silence ensues, which further delays the next item due. It is nonetheless important to note that, despite the occurrence of the silence, Kimi’s nonverbal conduct narrows down the possible semantic domain of the lexical element that will occur in the turn’s upcoming slot. This nonverbal conduct takes place in the following manner. First, having kept her gaze away from the recipient, Kimi blinks once (Figures 5.6 and 5.7) and gazes back to Yuko. Secondly, while shifting her gaze from an absent point to Yuko, Kimi produces lateral headshakes. Put differently, Kimi’s conduct establishes that the next item due is related to adding olives to certain meals, which is something that would make Kimi shake her head laterally. Kimi’s conduct makes it visible to the recipient that the upcoming component of her turn is associated with her emotional reaction to the addition of olives to food. At the beginning of line 6, Kimi gazes at Yuko (Figure 5.8), launching a mutual gaze, and produces the SSW *gyo*. What is significant here is that the SSW is followed by *tte kanji* ‘feel like,’ which explicitly confirms that the prior production of the sound *gyo* has rendered how she felt, thereby assessing the use of olives on meals. It should

be noted that no gesture accompanies the production of the SSW. Although gestures do not occur at this point, it has been described that Kimi has made facial gestures prior to the production of the SSW.

The recipient Yuko then repeats the sound symbolic unit that Kimi has produced with a slightly rising tone as *gyo(h)tte kanji(h)¿* 'do you feel like *gyo*?'(line 7), thereby initiating a repair (e.g., Schegloff, 1997). This prosodic design invokes Kimi's next turn, which reformulates what she has produced: *sonnani: (.) ireruka: mitaina* 'like, do you put that many olives on things?' (line 8). The production of this turn indicates Kimi's understanding that the prior production of the sound symbolic unit was the trouble source of the problem of understanding. There are three elements in Kimi's reformulated turn that indicate her emotional reaction cannot be clearly or concretely verbalized via non-SSWs. Frist, by prolonging the final sound *ni* in *sonnani:*, Kimi slightly delays the progress of the turn. Second, the following micro pause also delays the progression of the turn. Finally, the turn ending is marked with *mitaina* 'like,' which makes her prior elements indefinite. That way, Kimi designed the reformulated turn making it more accessible to Yuko, thereby working towards achieving intersubjectivity with Yuko.

The prior two segments have demonstrated that SSWs can be deployed when the speakers face trouble depicting immediate emotional stance with non-SSWs. In turn, these two segments have demonstrated that SSWs operate as a readily available resource to render nuanced emotional stances, which are not easy to depict with non-SSWs. In the next segment, the speaker utilizes various SSWs as resources to render a broad range of nuanced emotional stances that carry different levels of intensities.

In order to demonstrate speakers' deployment of SSWs that depicts diverse emotional stances, this study adopts the concept of "emotional scaling" (Prior, 2016). While the speaker

formulates or reformulates the ongoing turn, he/she incorporates the display of emotional stances, which will be presented graphically using emotional scales. The emotional scale is a visual representation of emotional intensity that differs in degree. Let us present a default scale in Figure 5.9 below. Each scale has minimization on one end and intensification on the other end in addition to zero in the middle.[3]

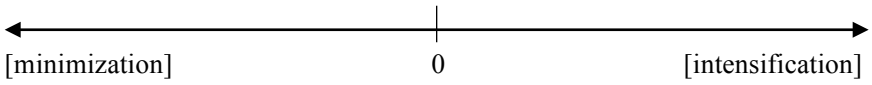


Figure 5.9. Default Emotional Scaling

Prior (2016) lists ten characteristics of emotional scaling, among which the following two points, related to emotional vectors, are most relevant to my study. One of them is about intensity.

> The location of a formulation along the scale is approximate, not exact. Therefore scales cannot be analyzed without examining the interactional and sequential environment of their production, their recipient design, and the relations among items (e.g., "upshot" and "angry" may be clustered together in one context and contrasted in another) (p. 215).

The other one is about the directionality.

> The items on the scale have directionality; that is, a series of formulations tends to either escalate or de-escalate along the scale. When they escalate, often within the activity of accounting or story production, they typically are associated with actions. This scaling is not unidirectional but may shift depending on the goals of the speaker and the actions represented (p. 215).

In order to visualize different degrees of emotional intensity displayed by each deployment of an SSW and the transitions between them, the following analysis will include presentations of emotional scales for each occurrence of an SSW as well as for the reformulated turn, consisting

[3] Prior's emotional scale is based on Labov's (1984) model on emotion and intensity.

of non-SSWs. By doing so, the analysis of the next segment shows SSWs' capability of instant depiction of emotional stances in face-to-face interaction.

The next segment comes from a talk among office colleagues named Yumi, Rina, and Aiko. They have been talking about whether or not they speak up when they have something to object to at work. Prior to the segment below, Rina claimed that she never spoke up. Even when Rina has something to say, she changes her mind on the next day and feels like *ma ikka* 'oh well whatever.' Responding to Rina's stance, Yumi confirms the opposite: Yumi and Aiko tend to speak up. The following analysis focuses on how Yumi delivers transitions of her emotional stances when she tries not to speak up by swallowing her frustrating anger. Yumi adopts four different SSWs—three unconventional SSWs and one conventional SSW— in depicting her emotional stances which are different in terms of how each stance arises and its intensity.

(5.4) [OL.1133][4]

37 Rina: *ah:↑demo sa: iou tto omowa hen kat tara=*
oh but FP say.will QT think NG PT if
'oh, but if you don't think you would speak up'

38 *=sono ba de pan te omou tte koto desho?*
that spot in SSW QT think QT thing TAG
'you would think like *pan* on the spot, huh?'

39 (0.4)

40 *˚pan te [ka˚ ˚˚*pi ( ) *˚˚*
SSW QT or
'*pa:n*, or *pi*( )'

41 Yumi: *[te ka demo gaman shite nomikon dara::*
QT or but tolerance do.and swallow when
'I mean, but when I put up with what I want to say and swallow it'

[4] This segment contains five instances of SSWs. However, *pan* and *pi* in line 40 are not targets of analysis for the following two reasons. First, these two SSWs are not produced by the main speaker, Yumi. Second, the turn in which these two SSWs occur turns out not to be responded by the main speaker.

Figure 5.10

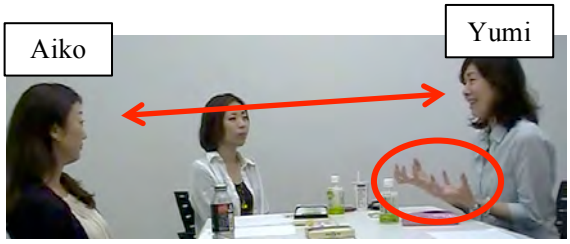


Figure 5.11

*

*

42 → *jiwajiwa:*(.) *uu:wa*[*::* *tte natte::*
gradually SSW QT become.and
'gradually (.) I become like *uu:wa::* and'

43 Aiko: [*u:::::::::::::::n*
'ye::::::::::::ah'

44 Rina: [°*un un un*°
'°uh huh uh huh uh huh°'

45 Yumi: [*ie toka de: nanka:,*
home etc. at like
'at home, like'

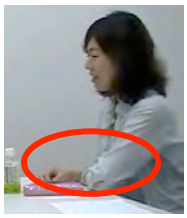


Figure 5.12

*

46 → *fu↓u tte kaetta ato ni=*
SSW QT went.home after when
'after I went home like *fu↓u*'

Figure 5.13

Figure 5.14

Figure 5.15

Figure 5.16

Figure 5.17

* * * *

47 → =>*UORA::*< *tte* [*natte:*
SSW QT become.and
'I go like >*UORA::*<, and'

48 Rina: [*u(h)*[*HAH HA(h) HA(h) HA(h) HA(h)*

49 Aiko: [uh ↑*HAH HAH HAH .H::*

Figure 5.18

*

50 Yumi: [↑*NANKA* [*yaPPAri* *OKASHI: mitai natte::*
like as.I.was.thinking wrong seem become.and
'I go like ↑something is WRO:ng as I was THINKING, and'

51 Rina: [*.HH:* [au-

Figure 5.21

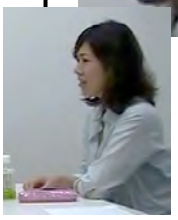


Figure 5.19 *Figure 5.20*

* *

52 Yumi: *˚sonna˚(.)yappa* *okashii ↑na ja nakute nanka*
such as.I.was.thinking wrong P CP NG.and like
'it's not such an attitude that something is ↑wrong as I was thinking, this is not the case, and like'

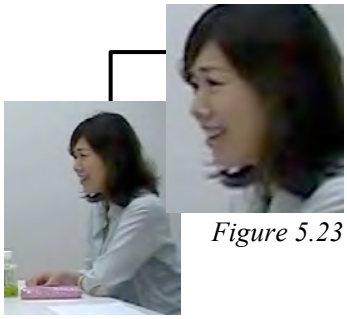


Figure 5.23

Figure 5.22

*

53 *a:↑::::tte nan nai?*
SSW QT become NG
'don't you go like *a:↑::::*?'

54 (0.4)

55 Yumi: *°so[re° iwana kattara*
that say.NG PT.if
'if you don't say what you want to say'

56 Aiko: *[atashi ↑kaette kara wasuren nen.*
I return.and after forget FP
'after I return home, I forget what I was annoyed at'

Rina asks Yumi a question, in lines 37 and 38. The question, however, does not receive any response from Yumi, as a 0.4-second silence occurs in line 39. Rina seems to reformulate her turn (line 40), but Yumi comes in and overlaps with Rina (line 41). At this point, Yumi initiates her turn with *teka* 'I mean,' which cancels the preceding question and delivers a situation *gaman shite nomikon dara* 'when I put up with what I want to say and swallow it' (line 40). Yumi's production of this situational background leads her to deliver a specific scene and a further depiction of her emotional stances through the deployment of SSWs that is accompanied by a hand gesture. Specifically, Yumi brings her right hand up, folding the fingers up and slightly shaking her hand horizontally. This hand movement is synchronized with the

conventional adverbial SSW *jiwajiwa* 'gradually'[5] (Figure 5.10). After a micro pause, Yumi vocalizes the non-conventional SSW *uu:wa::*, which is the first analysis target. During this vocalization, she gradually brings her hands up with the palms facing up (Figure 5.11). The SSW is followed by the quotative particle *tte* and the verb *naru* 'become,' which grammatically mark the prior SSW as the element rendering her emotional stance. The accompanying hand shape and movements iconically indicate that her emotional stance is something that has emerged and risen in the area of her chest. In this case, her emotional stance is frustrated rage, should we need to put a label on it. The following scale shows Yumi's stance rendered through the SSW *uu:wa::*.

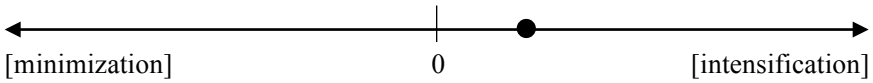


Figure 5.24. Intensity of Yumi's Emotional Stance Rendered through *uu:wa::*

As soon as Aiko hears the first part of the invented SSW *uu:wa::*, she provides a lengthened agreement token *u:::::::::::::::n* 'ye::::::::::::ah' (line 43). Rina's response through producing continuers slightly delayed and is delivered in a lower volume (line 44).

Rina's response overlaps with Yumi's continuation, which draws the specific scene in the setting, *ie toka de:* 'at home' (line 45) into the interaction. By producing the delaying item *nanka* 'like,' Yumi indicates she is having trouble continuing the turn. Yumi further describes the past experience by enacting herself at home, incorporating our second analysis target SSW: *fu↓u tte kaetta ato ni* 'after I went home like *fu↓u*' (line 46). The SSW *fu↓u* in this turn displays how she relaxes at home, as displayed by her bodily movements, placing her left elbow on the table (Figure 5.12). Moreover, the vocalization of *fu↓u* sounds lax. Yumi vocally and visually enacts

[5] This SSW is not included as a target of analysis, because it is an adverb that modifies the following verbal SSW, and is not directly associated with emotional display.

how she relaxes at home. The following scale shows Yumi's emotional stance rendered through *fu↓u.*

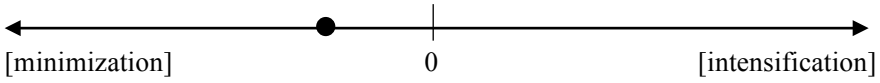


Figure 5.25. Intensity of Yumi's Emotional Stance Rendered through *fu↓u*

Yumi then brings her right hand with the palm facing upward in front of her body. Then she deploys the third analysis target SSW >*UORA::*< (line 47). There are two major multimodal features in the deployment of this SSW. First, the hand shape visually links the SSW >*UORA::*< with the prior turn, where she made a similar hand shape, the palm and fingers facing up (Figure 5.13). For reference, it was in line 42 where Yumi depicted, using her hands, that she gradually felt something rising in the area of her chest. In the same way she did while vocalizing the SSW *uu:wa::* (Figure 5.11), she raises her right hand (Figure 5.14). Secondly, as the transcription shows, the SSW is vocalized quickly, with the final vowel prolonged. Moreover, the SSW is produced in a louder volume and with an emphasized vocal quality. Here, the vocal gestural aspects of the SSW manifest themselves. Structure-wise, this SSW is followed by the quotative particle *tte* and *natte* 'become and,' indicating that the preceding SSW rendered how Yumi feels. Having coordinated these multimodal features,[6] Yumi renders, or even enacts her rage, which

[6] There is one more multimodal feature in Yumi's bodily conduct. Although Yumi depicts her frustrating rage, her face shows smiling. This facial display seems to be responsible for the occurrence of recipients' laughter. That is, the smiling face, which is considered "a milder" (Glenn, 2003:16) form of laughter, invites laughter from the recipient. A similar phenomenon, although not related to SSWs, is described by C. Goodwin (2007), who has demonstrated how various types of speakers co-exist within an utterance when the speaker reports past events (cf., original work on footing in Goffman, 1981). In his data, a speaker who is enacting her husband in a there-and-then situation produces laughter particles (i.e., transcribed as (h)). He contends that this laughter invites the recipient's laughter. In my database, there is another data segment in which the speaker displays a negative emotion while smiling, and subsequently the recipients laughs. Also, Ford and Fox (2010) have demonstrated that constructing a laughable is a multimodal phenomena in which vocal, facial, and body movements are all intertwined. Further detailed analysis is necessary to confirm the phenomenon in which the

she could not control and is rather aggressive and dynamic. The following scale shows Yumi's emotional stance rendered through the SSW >*UORA::*<.

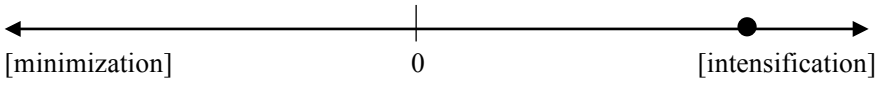


Figure 5.26. Yumi's Emotional Stance Rendered through >*UORA::*<

Though the production of the loud laughter (lines 48 and 49), the two recipients Aiko and Rina display their understandings that Yumi has just produced a laughable through the SSW. Yumi then reformulates the sound symbolic unit with the following non-sound symbolic unit: *NANKA yaPPAri OKASHI: mitai natte::* 'I go like ↑something is WRO:ng as I was THINKING, and.' This reformulated clause carries the following features. First, the reformulated clause shares the identical syntactic structure with the sound symbolic unit: *X natte* 'go like X' as indicated by boxes in lines 47 and 50 in the transcript (5.4). Secondly, both units include the visual display made through Yumi's hand gestures. When Yumi vocalizes *OKASHI:* 'WRO:ng,' she brings up her right hand with the palm facing upward, while shaking it (Figure 5.18). The shape and its movements are very similar to the ones that Yumi has made in line 47 while vocalizing the SSW >*UORA::*< (Figure 5.13). By designing the reformulated turn in this way, Yumi made her prior utterance more accessible to the recipients, which helps the recipients and herself achieve intersubjectivity. Moreover, this reformulation makes the core characteristic of SSWs even more visible. Specifically, Yumi rendered her emotional reaction through the SSW instantaneously, while in the reformulated turn, she needed and used non-SSWs. Both versions

---

speaker smiles while delivering troubles-telling leading to the speaker depicting emotional stances through SSWs and the subsequent laughter of the recipient.

involve prosodic stress as described, but the significant difference is that, in a short period of time, the SSW >*UORA::*< depicted her abrupt burst of indignation. In contrast, the production of the reformulated turn took longer than that of the single SSW. The following scale shows Yumi's emotional stance depicted through *NANKA yaPPAri OKASHI: mitai natte::* 'I go like ↑something is WRO:ng as I was THINKING, and'. The intensity can be considered about the same as the prior deployment of >*UORA::*<.

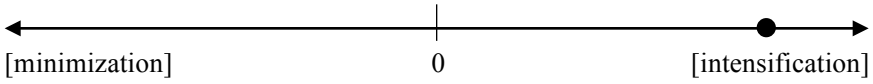


Figure 5.27. Yumi's Emotional Stance Rendered through *NANKA yaPPAri OKASHI: mitai natte:: 'I go like ↑something is WRO:ng as I was THINKING, and'*

As indicated by the use of the continuing form of *naru*, *natte* 'become and,' Yumi continues her turn. In line 52, she contrasts her emotional display against her prior depiction of rage. She tilts her head when she produces another emotional stance to contrast: *yappa okashii ↑na* 'something is ↑wrong as I was thinking' (Figure 5.19). Moreover, although the transcript cannot capture the vocal quality, Yumi vocalizes this unit in a slightly weak, gentle voice. Through the vocalization, Yumi manipulates the volume and voice quality, thereby depicting her emotional stance as if the emotional reaction is happening at the moment of interaction momentarily. Although Chapter 3 (SSWs as Vocal Gestures) has focused exclusively on SSWs, Yumi's vocal conduct in this reformulation discloses that a speaker is capable of depicting various emotional stances through manipulation of vocalization. The manipulation occurs not only when the speaker deploys SSWs, but also when he/she deploys non-SSWs. The following scale shows Yumi's emotional stance depicted through *yappa okashii ↑na* 'something is ↑wrong as I was thinking.'

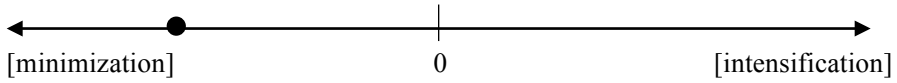


Figure 5.28. Yumi's Emotional Stance Rendered through yappa okashii ↑na *'something is ↑wrong as I was thinking'*

When Yumi grammatically negates the unit by adding *ja nakute nanka* 'is not the case, and like,' an element that projects contrast at the end of line 52, she makes a facial gesture which includes smiling lips and frowning eyebrows (Figures 5.20 and 5.21). This mixed facial gesture may display that her emotional display is in the transition phase from the gentle mode to the >*UORA::*< mode, or the frustrated-raging mode[7]. In line 53, she vocalizes the last analysis target SSW *a:*↑*::::*, which is marked by the quotative particle *tte* and the negative form of the verb *naru* 'become,' forming a question sentence to seek confirmation to Aiko as *a:*↑*::::tte nan nai?* 'don't you go like *a:*↑*::::*?' (line 53). The rising hand gestures do not occur this time. However, this instance of seeking confirmation is visually, or better said, facially, and vocally designed as another reformulation of the sound symbolic unit, which incorporated the SSW >*UORA::*<, in the following way. First, her vocalization *a:*↑*::::*? gradually goes up in pitch as indicated by the upper arrow in the middle. Although the prior vocalization of the SSW >*UORA::*< was not accompanied by such a rising pitch contour, Yumi gradually raised her right hand with the palm facing up. Also, in the same way as she vocalized the SSW >*UORA::*<, Yumi's face here displays a mixture of smiling mouth and frowning eyebrows (Figures 5.22 and 5.23). This reformulated unit shares these two parallel multimodal designs with the prior sound symbolic

[7] Alternatively, Yumi's smiling face can be analyzed as a way of showing the troubles-telling as non-serious one. Coupled with the previous segment, a further investigation will be needed regarding the deployment of smiling face during the delivery of troubles-telling.

unit. The following scale shows Yumi's emotional display depicted through *a:↑::::tte nan nai?* 'don't you go like *a:↑::::*?.'

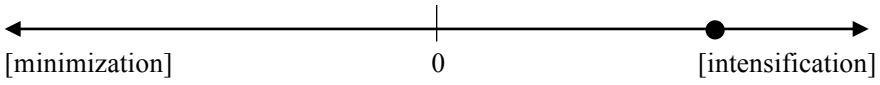


Figure 5.29. Yumi's Emotional Display Rendered through a:↑:::tte nan nai? *'don't you go like a:↑::::*?'

Although the end of Yumi's turn was designed to seek confirmation from Aiko, Aiko does not confirm what Yumi has depicted. (This part is not included in the transcript.) Aiko counters Yumi's suggestion by stating that she forgets about work once she returns home.

This section has demonstrated that emotional stances are not always straightforwardly depicted with non-SSWs. The analyses have revealed that, immediately after the speaker described a past experience, the deployed SSW served as a readily available resource to render his/her immediate emotional stance in the given situation. In particular, the first and second segments have shown that the participants' conduct indicates that the target emotional stances cannot be depicted through non-sound symbolic adjectives. The third segment has shown that the speaker utilizes SSWs when depicting his/her emotional stances with various intensities. By adopting Prior's (2016) concept of emotional scale, the analysis has visualized various intensities of emotional stances as rendered through the deployment of SSWs. By doing so, the analysis has revealed that SSWs are capable of instant depiction of emotional stances, carrying various intensities, by means of vocal gestures and accompanying bodily gestures. Thus, SSWs allow the speaker to show quick transitions of emotional stance one after another.

5.3 SSWs as a resource for the recipient to display understanding

This section demonstrates that not only the speaker, but also the recipient deploys SSWs when showing his/her understanding of the prior speaker's turn. Characteristically, the recipient produces the same, or a phonetically similar, SSW as the speaker does or has done. It is worth noting that the recipient who produces the SSW upon showing his/her understanding, in fact, has provided either a suggestion or question relevant to the specific experience the speaker had delivered. That is, the recipients had certain extent of knowledge or understanding of the speaker's troublesome situation. The recipient's suggestion or question then prompts the speaker to provide an account of the specific experience. As the prior section has demonstrated, the speaker's delivery of experience guides him/her to depict their emotional reaction to the given situation through the deployment of SSWs. The analysis in this section concentrates on how the recipients monitor the development of the ongoing turn, wherein the speaker constructs the turn toward SSW deployment that depicts his/her emotional stance.

In the first segment, the recipient provides a candidate item through an SSW, which is phonetically similar to the one produced by the speaker. In the second segment, the recipient chorally produces an SSW with the speaker (Lerner, 1996, 2002). The recipients' conduct, producing the same or phonetically similar SSWs, apparently contradicts with what the previous section has established: the speaker deploys SSWs to convey his/her nuanced emotional stance, and thus the recipients could be uncertain about the meaning. However, as has been mentioned above, the speaker's turns that incorporate the SSW in this section are prompted by the recipient's suggestion or inquiry about a certain situation the speaker is in. That is, the recipients possess relevant knowledge to some extent. The knowledge about the certain situation, which the

speaker has experienced, makes it possible for the recipients to show their understandings through SSWs in the similar or the same way the speaker does.

Furthermore, the SSWs that the recipient produces are not totally new invented SSWs. In the first segment, the target SSW is a conventional SSW. In the second segment, the target SSW is not a conventional one, but is initially produced solely by the speaker whereby the speaker established the target SSW as a shared one. These two segments reveal that the way the turn unfolds allow the recipient to anticipate what sort of sound symbolic emotional depiction comes next. In other words, an SSW that carries a shared understanding can be produced by the recipients, who do not possess firsthand experience in a certain situation. Through the production of the shared SSW, the recipient displays that the speaker and the recipient himself/herself have achieved intersubjectivity.

The first segment takes place among office colleagues: Aiko, Rina, and Yumi. In the same way as the previous segment (*gyo tte kanji* 'I feel like *gyo*' in (5.3) in Section 5.2), the speaker here also indicates that she has trouble with continuing the turn. Specifically, the speaker produces a pause, a delaying item, and facial gestures, and then produces the SSW that renders her emotional stance. That is, the teller deploys the SSW as a resource to render the emotional stance that she cannot verbalize right away.

The participants have been talking about Aiko's right wrist, which has suffered from *kenshooen* 'tenosynovitis.' Prior to the segment, Aiko has informed the recipients that she had a sense of discomfort in the wrist that she did not like, which solicited Yumi's sympathy to Aiko's situation at work. Responding to Yumi, Aiko explained how she handled the situation: she used fingers that she normally does not use.

In the following segment, Yumi suggests using a different arm (line 18). Put differently, Yumi was able to speculate that Aiko could use a different arm to avoid the sense of discomfort in the wrist she did not like. Aiko then responds *ah yatten kedo* 'oh I did it but' (line 19), claiming access to the situation as her firsthand experience. The following analysis focuses on how Aiko continues the turn in which she delivers an explanation of her experience. It also focuses on how the explanation prompts the forthcoming depiction of her immediate emotional stance through the SSW. The analysis also highlights how the recipient—in particular Yumi—monitors the ongoing turn and produces the SSW to display her understanding of Aiko's situation. Here, Aiko invents the SSW *i*, whereas Yumi produces the conventional SSW *ira*.[8]

(5.5) [OL.2039]

18 Yumi: *kocchi gawa mo renshuu shitoku tte no wa?*
this side too practice do.have.done QT N TP
'how about practicing with this hand too?'

19 Rina: [*a, >nanka<*
oh like
'oh, like'

20 Aiko: [*a, yatten kedo::*
oh did.FP but
'oh I did that but'

21 Rina: *un*
'uh huh'

22 Aiko: *dekin nen settee mo dekin nen kedo::,*
can FP configuration too can FP but
'It can be done, you can set it that way but'

[8] The SSW *ira* was added to an ordinary dictionary in 2014, which is relatively recent, according to the 1960th issues of *Anan* magazine (June 2015), whereas *i* is not found in any dictionary. *Anan* is a fashion magazine that targets women mainly in their 20s and 30s. This issue deals with recent trends and uses of Japanese language.

23 Rina: *un.*
‘uh huh’

Figure 5.30 *Figure 5.31*

* *

24 (0.8)

Figure 5.32

Figure 5.33

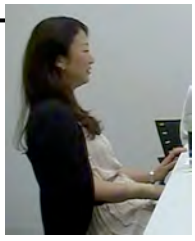


Figure 5.34

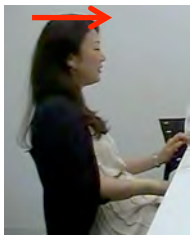


Figure 5.35

* * *

25 Aiko: → °*ano*° (0.4) [*i tte natte,*
uhm SSW QT become.and
‘uhm (0.4) *i* is how I felt, and’ ((keeping the Japanese word order))
‘uhm (0.4) I felt like *i,* and’ ((more natural word order in English))

26 Yumi: → [*ira tto suru?*
SSW QT do
‘*ira* is how you feel?’ ((keeping the Japanese word order))
‘do you feel like *ira*?’ ((more natural word order in English))

27 Aiko: → *ira tte natte*
SSW QT become.and
“I felt like *ira* and”

In response to Yumi's suggestion, Aiko reports that she tried what Yumi suggested: *ah yatten* 'oh I did it' (line 19), which is however marked by *kedo::* 'bu::t.' The ending vowel is prolonged, which delays the progression of the turn. From this point onward, Aiko delivers what she has experienced. Aiko continues her turn by informing the recipients of the fact that what Yumi has suggested was actually possible: *dekin nen settee mo dekin nen* 'it can be done, you can set it that way' (line 22). The element *nen* that Aiko deployed twice in this turn is the equivalent of the so-called standard Japanese *noda*, which indicates that the speaker is in an explanatory mode about what she has just said (Kuno, 1973). However, Aiko again produces the prolonged *kedo::* 'bu::t' at the end of the turn, displaying that Aiko is going to continue her turn, which contradicts what she has delivered up to this point.

After Rina's production of the continuer *un* 'uh huh' (line 23), a 0.8-second silence takes place (line 24). During this silence, Aiko visibly depicts her emotional stance by deploying head movements and facial gestures. As Figures 5.30 and 5.31 show, Aiko laterally shakes her head. At the end of the head shake, Aiko quickly blinks twice. She then produces the delaying item ˚*ano*˚ ''uhm˚'' in a smaller voice (line 25), while she holds a mutual gaze with Aiko (Figure 5.32). This mutual gaze solicits Yumi's response later in line 26. In line 25, another silence follows, which lasts for 0.4 seconds. During this silence, Aiko closes her eyes and knits her eyebrows together (Figure 5.33).

At this point, Aiko enacts her emotional stance through the sound symbolic unit *i tte natte* 'I felt like *i*, and.' In vocalizing the sound symbolic unit, Aiko lays the stress on *i*, as if she was moaning in agonizing pain. Furthermore, her head quickly comes forward at this very moment (Figures 5.34 and 5.35). The SSW is immediately followed by the quotative particles *tte* and the

continuing verb form *natte* 'become and,' both of which grammatically indicate that the preceding SSW has rendered her emotional stance in the given situation.

At the very moment at which Aiko renders her emotional stance through the SSW *i*, the recipient Yumi simultaneously offers a candidate item of Aiko's emotional stance through the SSW *ira*. This candidate item is also grammatically marked as an emotional stance through the deployment of the quotative particle *tte* and the verb *suru* 'do' (line 26). The original speaker Aiko then confirms Yumi's version of the SSW (i.e., *ira*) by repeating it with a different verb attached: *ira tte natte* 'I felt like *ira*, and' (line 27). Yumi's conduct demonstrates the following two points. First, Yumi's production of the candidate item reveals that Aiko's turn, in which she delivered her experience, was transparent. That is, Aiko's turn had provided a clue to Yumi that an element depicting Aiko's emotional stance was forthcoming. Accordingly, Yumi was able to provide the candidate item (i.e., *ira*), which was then confirmed by Aiko, the original speaker. Aiko's confirmation, at the same time, provided a reformulation of her original SSW. Second, the speaker who actually experienced an event might be exclusively eligible to produce unconventional SSWs when displaying his/her emotional stance. As mentioned, Aiko's version *i* is less conventionalized than Yumi's version *ira*. Having had no direct experience, the recipient is not in a position to produce an unconventional SSW that appears to be directly linked with the speaker's firsthand experience.

The next segment also demonstrates how the recipient shows his/her understanding of the speaker's, or teller's, troublesome events through an SSW. The segment takes place among three close friends, Chie, Kana, and Mako, who have been talking about people and their behaviors that frustrate them at work. Chie's troubles-telling is prompted by Mako's inquiry about Chie's situation at work: *sono ko no sagyoo de ashihipparare tari suru?* 'do the assistant's tasks cause

you trouble?.' By producing this question, Mako displays that she is aware of the work situation Chie is in. Responding to Mako's question, Chie first produces *n:* 'yeah,' during which she nods. Starting at this point, Chie's complaints, specifically about the assistant, go on. For example, the assistant asks Chie's confirmation even for sending a sheet of paper by mail or by fax. As a result, Chie has to stop her hands and check everything the assistant has done. Once Chie's list of complaining comes to a break, Mako compliments her explained effort as *era* 'awesome,' and the other recipient Kana displays her sympathy by saying *sore wa shindoi na* 'that's so tough.'

The first part (5.6) below shows that Chie introduces other coworkers who cause even more trouble. It should be noted that these coworkers were already mentioned prior to the summarized scene above. My analysis first focuses on how Chie's turn unfolds, while she deploys hand placements and movements that visually depict the relationship between the coworkers and Chie. The analysis then focuses on how the depiction of the relationship prompts Chie to deploy the SSW through which she enacts her emotional reaction to that particular situation. The second part (5.7) confirms that the story recipient Mako, who had asked the question that prompted Chie to tell the story, understands Chie's emotional stance by chorally producing the same SSW that Chie produced.

(5.6) [KK.2028-2]

Figure 5.36

★

45 Chie: *=honde sono namaiki no nijuugo to=*
and that impertinent LK twenty.five and
'and that impertinent twenty five years old one and'

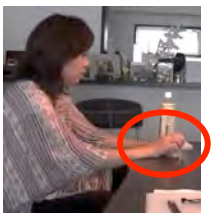


Figure 5.37

★

46 *=hanashi tsuujihin onaidoshi no hito wa:*
talk get.through same.age LK person TP
'the person my age who doesn't understand what I say'

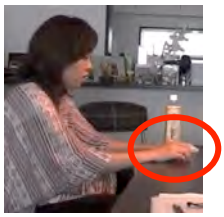


Figure 5.38

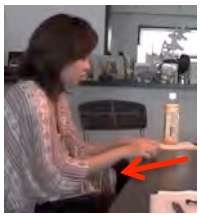


Figure 5.39

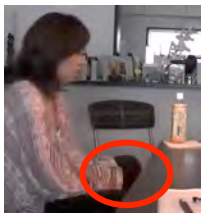


Figure 5.40

★ ★ ★

47 *mattaku seehen kara[: zenbu Chie ni kun nen*
at.all don't.do thus all Chie to come FP
'don't work at all, and so all the work is falling to me'

48 Mako: *[nu ↑hu hu hu hu hu hu*

49 *hu [hu hu*

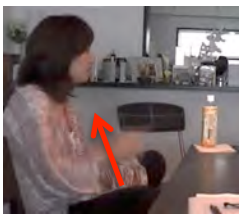


Figure 5.41

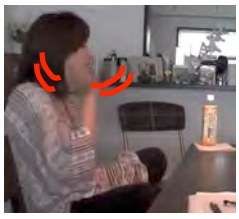


Figure 5.42

* *

50 Chie: → (.)[>*UUUn*< *tte nan* [*noni*
SSW QT become and.further
'I become like >*UUUn*< and further'

51 Kana: [°*n:::*°
'°uh huh:::°'

52 Mako: [*h.hh*

Chie introduces two coworkers she talked about prior to the segment: *sono namaiki no nijuugo* 'that impertinent twenty-five-year-old one' (line 45) and *hanashi tsuujihin onaidoshi no hito* 'the person my age who doesn't understand what I say' (line 46). Chie uses her hands to depict the relationships between the two coworkers and herself. When introducing the former, Chie places her left palm facing down on the table (Figure 5.36). When introducing the latter, Chie does the same using her right palm (Figure 5.37). In line 47, Chie explains that those two coworkers *mattaku seehen* 'do not work at all,' which is marked with the conjunctive particle *kara* 'thus,' indicating that she is going to deliver the consequence of having two coworkers who do not work. At this point, Chie's hands stay in the position in which she has placed them already on the table (Figure 5.38).

The indicated consequence of having two coworkers, who do not work, is delivered next. Chie's hand movements render a visible connection between the prior components and the

forthcoming components of the turn. As Chie continues her turn: *zenbu Chie ni kun nen* 'all the work is falling to me' (line 47), she brings her hands close to her body (Figure 5.39). In the same way as the teller in the previous segment did, Chie marks the end of delivering the situation with the element *nen*, which indicates that she is in the explanatory mode regarding what she has said (Kuno, 1973). When the turn comes to an end, her hands are placed on her lap (Figure 5.40), which visibly marks the end of the explanation of the scene at work. While Chie produces that turn, Mako produces laughter. While Mako's laughter continues, Chie overlaps with Mako and enacts her emotional stance through the SSW >*UUUn*<, which is marked by the quotative particle *tte* and the conjugated verb *nan* 'become' (line 50), indicating that the preceding SSW rendered Chie's emotional stance. That is, the emotional stance is displayed immediately after Chie finished explaining her troublesome situation at work. This SSW is vocalized in a voice louder than the surrounding talk and in an emphasized manner. Moreover, the accompanying hand movements are rhythmically synchronized with the vocalization of the SSW. Immediately before the vocalization of the SSW, Chie's hands start moving upward (Figure 5.41). The hands then reach the head level, where she slightly shakes her hands in sync with the vocalization of >*UUUn*< (Figure 5.42). These various semiotic resources are all integrated into one, rendering Chie's emotional stance in that particular scene. The sound symbolic unit is then followed by *noni* 'and further,' which indicates Chie is going to continue her telling. Both recipients have been gazing at Chie, displaying their stance as the listeners. At the same time as Chie vocalizes the SSW, Kana simply acknowledges Chie's situation by producing °*n:::*° '°uh huh:::°' (line 51). On the other hand, as soon as Mako hears the verb in the sound symbolic unit, she produces an audible inhale, although only slightly.

In the second part of the segment, which is shown below, Chie continues delivering telling and renders her emotional stance. Again, she places her hands on the table to depict the human relationships in the workplace. This account of experience again prompts Chie to depict her emotional stance through the SSW. This time, the recipient Mako also chorally produces the same SSW as Chie, claiming her understanding of Chie's situation which had caused her to take the emotional stance rendered through >*UUUn*<.

(5.7) [KK.2028-3]

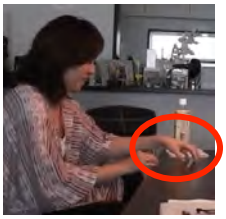


Figure 5.43

53 Chie: *.hh [koitsu kara mo=*
this.person from too
'.hh from this person too'

54 Mako: *[.hh*

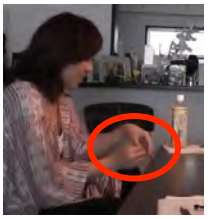


Figure 5.44

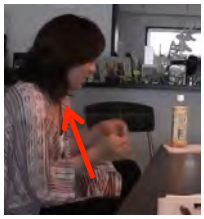


Figure 5.45

Figure 5.46

55 Chie: *=nanka iyana koto toka iwarete*
something unpleasant things etc being.said.and
'I hear something unpleasant, and so'

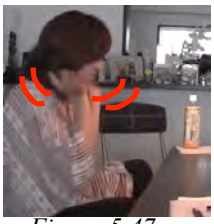


Figure 5.47

*

56 → [>*UUUn*< *tte nan nen*
SSW QT become FP
'I become like *UUUn*'

57 Mako: → [>*UUUn*< *tte natte*
SSW QT become.and
'you become like *UUUn*, and'

58 Kana: °*n:::::* °
'°uh hu:::h°'

In line 53, Chie draws another person into the interaction by deploying *koitsu kara mo* 'from this person too,' during which she places her left palm facing down on the table (Figure 5.43). As the deployment of *mo* 'too' indicates, Chie continues explaining that this person also causes trouble: *nanka iyana koto toka iwarete* 'I hear something unpleasant, and so' (line 55). While verbalizing this turn, Chie deploys hand movements that indicate how the upcoming course of the turn will progress. When she produces *iyana* 'unpleasant,' she brings her hands close to her body (Figure 5.44). The following noun phrase *koto toka* 'things etc.' co-occurs with the hand-raising movement (Figure 5.45). When Chie produces the verb *iwarete* 'being said, and,' her hands reach to her head level (Figure 5.46). At this point, Chie's way of placing her hands next to her head is similar to her previous movements. It was in line 50 when she first deployed the SSW >*UUUn*< (Figure 5.42) for rendering her emotional reaction at the scene of the event. Over the course of the turn, the recipient Moko observe Chie's delivery of the turn with a

constant gaze. The visual correlation between the two hand movements subsequently prompt Mako to produce the same SSW as Chie is about to do. In line 56, as visually projected, Chie enacts her emotional stance once more through the SSW >*UUUn*<, followed by the quotative particle *tte* and the verb *naru* 'become.' As Figure 5.47 shows, Chie shakes her hands next to her head. Here, the recipient Mako overlaps with Chie by producing the same SSW >*UUUn*<, displaying her understanding that Chie's situation caused this particular emotional reaction. The vocalization of the two SSWs are entirely in sync with one another. That is, the way Chie's turn unfolds makes the upcoming slot wherein she depicts her emotional stance recognizable for Mako, thus allowing Mako to chorally produce the same SSW.

This section has examined cases in which the recipient displayed his/her understanding of the prior speaker's/teller's emotional stance, whose display had been prompted by the speaker's own delivery of the troublesome situation at work. The speaker delivered the troublesome situation as an account in the answer to a suggestion or inquiry, which had been provided by the recipient. The analyses revealed that the recipient who made a suggestion or inquiry already had an understanding of the speaker's/teller's troublesome situation to the extent that they were able to make the suggestion or the inquiry. That is, when possessing relevant knowledge about the speaker's troublesome situation in advance, the recipient is eligible to deploy SSWs as a way of claiming an understanding of the prior turn where the speaker rendered their emotional stance. The recipient might have gained the relevant knowledge while the prior interaction advanced, or knew it even before their interaction.

5.4 Summary of chapter 5

One of the goals of this dissertation is to clarify what SSWs can accomplish in face-to-face interaction. To pursue this goal, this chapter focused on cases in which the speaker deployed SSWs in depicting his/her emotional stances, and demonstrated how SSWs rendered the speaker's immediate emotional reaction to the given situation. Section 5.2 demonstrated how SSWs served as a readily available resource to depict one's emotional stances. The analyses showed that the speaker provided explanations of a specific scene, which led the speaker to depict his/her emotional stances through the deployment of SSWs. The analyses also showed that speakers' conduct revealed the target emotional stance could not be easily displayed through non-SSWs. Section 5.3 revealed that once an SSW was deployed by a speaker, who rendered their emotional reaction to a given situation, recipients could use this newly introduced resource to display their understanding of the prior speaker's stance. The recipient, who had provided the question or suggestion, claimed their understanding of the teller's emotional stance by producing the same or a phonetically similar SSW. By doing so, the recipient audibly and visually displayed that he/she and the speaker achieved intersubjectivity.

Through the examination of the deployments of SSWs in different sequential environments, this chapter also disclosed a distribution tendency between conventional SSWs and unconventional SSWs. More precisely, unconventional invented SSWs could be produced exclusively by the speakers who had firsthand knowledge of a given situation. This phenomenon shows that the display of emotional stance is directly linked with the experiences which one has had in a certain situation. In contrast, conventional, or established, SSWs can be produced by either the speaker or the recipients. However, it does not mean that conventional SSWs does not occur when the speaker renders his/her emotional stance. As demonstrated in this chapter,

speakers can deploy both conventional and invented SSWs when rendering their emotional stances.

Part 4

Conclusion

Chapter 6

Conclusion

6.1 Contributions of this study

As summarized in the Introduction, many previous studies on SSWs paid attention to semantics of SSWs by focusing on the meaning of each SSW. This study has taken a different approach toward SSWs, while examining SSWs as an interactional resource when one is participating in ongoing face-to-face interaction. This study is the first to offer analyses of SSWs in naturally occurring conversation among native speakers of Japanese.

Building on the previous findings about SSWs' capability of depictions, Part 2 (Chapters 2 and 3) demonstrated that speakers make use of hand gestures and vocal gestural aspects of SSWs upon vocalization of SSWs. In analyzing the data, this study classified SSWs into three groups, while previous studies treated SSWs as a word class. The first contribution of this study is that, by categorizing SSWs in this way, it offered a new perspective towards our understanding of SSWs. For example, this study has revealed that invented SSWs are pervasive in naturally-occurring everyday conversation. Also, the numerical data, presented in Chapter 2, illustrated the differences in frequency of bodily gestures (i.e., hand gestures, facial gestures, and body gestures) between conventional, creative, and invented SSWs. The use of hand gestures had already been reported in previous studies (e.g., Kita, 1997), but the current study is significantly different from the previous studies in the following respects. First, this study revealed that the degree of conventionality seems to be related to the gesture occurrences. Specifically, the data showed that while less than fifty percent of conventional SSWs were accompanied by hand gestures, approximately eighty percent of creative and invented SSWs occurred along with hand

gestures. Second, Chapter 2 demonstrated that gestures visually represent the speech content and contribute to embodying the meaning of the SSW in interaction. By examining the conversation segment that contained two occurrences of the same SSW, the analysis demonstrated that, although the first occurrence of the SSW accompanied a hand gesture, the second occurrence of the same SSW produced by a different speaker did not accompany a hand gesture. The analysis revealed that the absence of a hand gesture in the second appearance of the SSW indicated the participants' understanding that the meaning of the SSW had already been established in the interaction. For that reason, the hand gestures that could serve as a meaning compensation did not occur. The clear categorization of SSWs made these findings possible.

As mentioned, SSWs were considered "vocal gestures." However, there were no studies describing SSWs as vocal gestures, so this study, for the first time, disclosed how speakers manipulate the phonetic shape of the word and vocalize the word when depicting its respective referent. The second contribution of this study is that it has revealed that vocalization is an inseparable aspect of the deployment of SSWs. Chapter 3 first synthesized how previous studies explained the meaning of *pon*, taking sound symbolism and iconicity into consideration. It then provided analyses of several instances of SSWs, which included occurrences of one or more *pon*. The analyses revealed that the sound chunk *pon* can be vocalized in various ways, such as *po:ːn*, >*ponponponpon*<, or <*POnPOnPOn*>. In face-to-face interaction, these various phonetic shapes depicting something abstract, such as motion, invisible concepts, and motionless objects, worked together with the speaker's bodily conduct. The referents of the SSWs were all abstract in a sense that they have no physical or concrete existence.

Furthermore, in relation to the concept of iconicity, my analyses suggested that when it comes to actual use of SSWs, at least one of the co-occurring multimodal features (i.e., hand

gestures, facial gestures, body gestures, and/or vocal gestures) displays an iconic aspect that depicts the referent visually and/or audibly. This finding is in line with Dingemanse's (2011) claim that not all SSWs ("ideophones" in his terminology) are iconic, focusing on a word form and its meaning. I expanded it further by examining the speakers' audio and visual display that work on an iconic level.

The primary concern of Chapter 3 was how SSWs operate as vocal gestures. However, the analyses turned out to shed light on another intriguing issue surrounding the use of SSW—particularly, the synchronization of hand gestures upon vocalization of the SSWs. The third contribution of this study is that it revealed an order related to the co-occurring hand gestures. Specifically, the analyses demonstrated that when the target of the SSW depiction was clear to the recipients, hand gestures were synchronized with the sound beat of the SSW. In contrast, when the target of the SSW depiction was still unclear to the recipients, hand gestures occurred but were not precisely coordinated with the sound beat of the SSW. In one of the cases, this happened when the speaker engaged in describing a scene in a telling activity. In another case, this happened when the speaker gave a piece of advice to the recipient. Put differently, gesture occurrences are shaped by various factors—most importantly, how the speaker constructs the ongoing turn. Thus, the key perspective the analysts need to consider is how ongoing turns that incorporate SSWs unfold. As noted, previous studies, such as those by Kita (1993), found that mimetics are "almost always" (p. 82) accompanied by a hand gesture that synchronizes with speech. My data analyses highlighted an area of this field which has not, up to the present time, been investigated yet regarding gesture use upon deployment of SSWs.

In the Introduction, I pointed out that the former characterizations of SSWs, such as concrete and abstract, are inconsistent with each other. The forth contribution of this study is that

it has revealed that meaning of SSWs can be either concrete or abstract depending on which point of the interaction the analyst pays attention to. In my analyses, both of these apparently inconsistent characteristics appeared to have emerged during the process in which the speakers' turns unfold. In particular, the analyses in Section 4.3 demonstrated how these semantic characteristics emerged while analyzing the data that revealed the characteristics of possibility and limitation when deploying SSWs. Without thorough examination of moment-by-moment progression of interaction, the inconsistency would not be explained.

As noted in the Introduction, SSWs are recognized as "quintessentially social" (Childs, 1994:63) linguistic items and "the closest linguistic substitute for a non-verbal, physical act" (Kunene, 2001:183). The fifth contribution of this study is that it analyzed SSWs in naturally occurring interaction and revealed how SSWs work as linguistic items in social interaction. Building upon the findings about gesture occurrences and SSWs serving as vocal gestures, Part 3 (Chapters 4 and 5) demonstrated what the speaker can accomplish by deploying SSWs in interaction. In particular, Chapter 4 first demonstrated how the speaker concretizes absent objects through vocalization of SSWs. The analyses clarified the series of steps of this practice: the speaker refers to the object by using the demonstrative pronoun *sore* 'that one,' and then depicts the property of the object through deployment of SSWs. However, there were cases that indicated SSWs' possibilities and limitations in depicting the referents. That is, SSWs were not always capable of depicting attributes of the referent when concretizing the objects. The analyses demonstrated that when the existence of the object had not been concretely established, the subsequent depiction of the property called for a reformulation. In contrast, when the existence of the object had been made clear to the recipients, the subsequent depiction of the property through deployment of SSWs accomplished the concretization of the object. Chapter 4 then

concentrated on cases in which the speakers make use of this practice of concretizing an object through an SSW when facing trouble with finding a word and engaging in a word search activity. The analyses showed that the deployment of the SSWs allowed the speaker to manage the situation. The speaker instantly concretized an object that did not exist in the interactional space through the deployment of the SSW, thereby making the interaction forward.

Chapter 5 demonstrated that SSWs work as a readily available resources for the speaker to depict his/her emotional stance. The first half of the chapter revealed that speakers deployed SSWs as a resource to render his/her immediate emotional stance that was not easy to depict by means of already existing non-SSWs in a way the recipient understood. The chapter also demonstrated that SSWs can work as a resource for the recipient to show his/her understanding of the previous speaker's emotional stance. The second half of the chapter focused on cases in which the recipient produced an SSW, thereby demonstrating their understanding of the prior turn in which an SSW was incorporated that rendered the speaker's emotional stances. The analyses revealed that the speaker's turn design provided the recipient with a clue that a depiction of emotional stance comes next, and more precisely, that emotional depiction through SSWs will follow subsequently. In turn, these two patterns uncovered sequential environments wherein conventional and non-conventional SSWs could occur respectively. Namely, conventional or shared SSWs occur when the recipients demonstrated their understandings of the prior turn, while non-conventional or brand-new SSWs occur when the speaker depicted his/her immediate emotional stances.

6.2 Implications and possible future directions

This final section discusses implications and how this study could be expanded in the future.

First, this study revealed what speakers can do through deployment of SSWs. Future studies are expected to identify different interactional achievements made through deployment of SSWs that bear descriptive force. Further examination of data from everyday conversation would identify different achievements through deployment of SSWs. Furthermore, cross-linguistic analysis of the phenomena found in this study would expand our understandings of the nature of SSWs. Some scholars claimed that Japanese is one of the languages that is rich in SSWs, while Indo-European languages do not have many SSWs. However, Hida and Asada (2002) pointed out that European languages also have sound expressions, although in written form, which are created by designing letters in various ways, for example, in *manga* 'comics.' Furthermore, from my personal experience, native speakers of English also use some kind of sound expressions that I could consider as SSWs in daily English conversation (e.g., I went there like *psh:::*). In order to have a better understanding of SSWs, scholars should not limit the target languages to Asian and African languages, but be open to any languages in the world.

Second, the analyses have delineated the possible sequential positions where the sound beat of the SSWs synchronizes or does not synchronize with the co-occurring hand gestures. Several previous studies on synchrony of speech and gesture took a quantitative approach, without paying attentions to how interaction unfolds. With more data, future research is needed to confirm the order of gesture occurrences and gesture-beat synchronization upon the vocalization of SSWs. By taking the participants' viewpoints in analyzing naturally-occurring interaction data, analysts are expected to follow why the speaker depicts something using SSWs at that moment of interaction. To make the research fruitful, the concepts such as McNeill's gesture phases (i.e., preparation, stroke, and retraction) should be integrated with the methodology of Conversation Analysis that focuses attention to "why that now" (Schegloff &

Sacks, 1973:299) from the viewpoint of participants in social interaction.

There are three more remaining issues that need to be addressed with respect to SSWs studies. First, Japanese SSWs are not only used in everyday conversation but also in some institutional interaction, such as in medical settings (Iwasaki, Vinson & Vigliocco, 2007; Miura & McGloin, 2008). As far as I know, SSWs in such institutional settings have not been studied yet. Thorough examination of interaction in various institutional settings may reveal different aspects of meaning construction process of SSWs as well as the interactional achievements that participants achieve.

Second, having examined the unfolding process of interaction, this study sheds light on the issue of sematic mapping. As introduced in the Introduction, while summarizing research history of SSWs, Akita (2013b) pointed out that the mechanism of SSWs' semantic mapping had not been clarified yet. This study casts doubt on whether the mechanism of semantic mapping can be explained by examining a single SSW detached from the linguistic and social context. In the naturally occurring interaction examined in this study, we observed that the speaker and the recipients went through the process of achieving intersubjectivity: for example by reformulating the turn in response to the repair initiation produced by the recipient. Put differently, participants displayed that SSWs do not necessarily carry inherent meaning that is shared among the participants. Instead, their understanding of the SSWs was gradually constructed through the ongoing face-to-face interaction.

Third, although this study did not include conversation among Japanese language learners, there are some issues that need attention. As Iwasaki, Vinson, and Viglioccoc (2007) summarized, even advanced learners of Japanese face trouble with acquiring SSWs. As long as SSWs are pervasive in everyday conversations, learners need, at least, to be able to perceive

what SSWs depict. The analyses have shown that SSWs co-occur with various multimodal clues that provide what SSWs refer to. Given this observation, SSWs may not necessarily be challenging items for learners in conversational context. Examination of interaction among native and non-native speakers may reveal how non-native speakers deal with the linguistic elements that have been considered difficult to understand.

As a final note, I want to raise another issue of whether or not the target item of this study is words. This study has called the target objects sound symbolic words; however, the analyses have demonstrated that their characteristics in face-to-face interaction seem to go beyond the definition of words. In the field of linguistics, there are various definitions of words depending on the linguistic unit focused on. For example, Bussman (1996) stated that, as one of the definitions of words, words "can be described syntactically as the smallest permutable and substitutable units of a sentence" (p. 512). If we put aside the issue of whether or not an utterance can be considered a sentence, SSWs fall into the category of words. However, semantically speaking, Crystal (2008) explained that a word is "[a] unit of expression which has universal intuitive recognition by native-speakers, in both spoken and written language" (p. 521). The analyses have shown that participants went through the process of understanding each other's turns. Thus, Crystal's semantic explanation cannot be applied to SSWs, especially to invented SSWs. There is room for further consideration of how to name what this study has called SSWs.

References

Akita, K. (2012). Toward a frame-semantic definition of sound-symbolic words: A collocational analysis of Japanese mimetics. *Cognitive Linguistics, 23*(1), 67-90.

Akita, K. (2013a). *Kyouki tokusei kara miru onomatope no fureemu imiron* [Frame semantics of onomatopoeia]. In K. Shinohara & R. Uno (Eds.), *Onomatope kenkyuu no shatei: Chikazuku oto to imi* [Sound symbolism and mimetics: Rethinking the relationship between sound and meaning in language] (pp. 101-116). Tokyo: Hitsuji Shobo.

Akita, K. (2013b). *Onomatope, otoshoochoo no kenkyuushi* [Research history of onomatopoeia and sound symbolism]. In K. Shinohara & R. Uno (Eds.), *Onomatope kenkyuu no shatei: chikazuku oto to imi* [Sound symbolism and mimetics: Rethinking the relationship between sound and meaning in language] (pp. 333-364). Tokyo: Hitsuji Shobo.

Alibali, M. W., Flevares, L., & Goldin-Meadow, S. (1997). Assessing knowledge conveyed in gesture: Do teachers have the upper hand?. *Journal of Educational Psychology, 89*, 183-193.

Amha, A. (2001). Ideophones and compound verbs in Wolaitta. In F. K. E. Voelts & C. Kilian-Hatz (Eds.), *Ideophones*. Amsterdam/Philadelphia: John Benjamins.

Antaki, C. (2008). Formulations in psychotherapy. In A. Peräkylä, C. Antaki, S. Vehviläinen, & I. ILeudar (Eds.), *Conversation Analysis of psychotherapy* (pp. 26-42). Cambridge: Cambridge University Press.

Antaki, C., & Widdicombe, S. (1998a). *Identities in talk* London: Sage.

Antaki, C., & Widdicombe, S. (1998b). Identity as an achievement and as a tool. In C. Antaki & S. Widdicombe (Eds.), *Identity in talk* (pp. 1-14). London: Sage.

Atkinson, M., & Heritage, J. (1984). *Structures of social action*. Cambridge: Cambridge University Press.

Bartens, A. (2000). *Ideophones and sound symbolism in Atlantic creoles*. Helsinki: Finnish Academy of Sciences and Letters.

Bilmes, J. (2011). Occasioned semantics: A systematic approach to meaning in talk. *Human Studies, 34*(2), 129-153.

Burrows, D. (1990). *Sound, speech, and music*. Massachusetts: The University of Massachusetts Press.

Bussman, H. (1996). *Dictionary of language and linguistics*. London: Routledge.

Chafe, W. (1985). Linguistic differences produced by differences between speaking and writing. In D. R. Olson, N. Torrance, & A. Hildyard (Eds.), *Literacy, language, and learning* (pp. 105-123). Cambridge: Cambridge University Press.

Childs, G. T. (1994). African ideophones. In L. Hinton, J. Nichols, & J. J. Ohala (Eds.), *Sound symbolism* (pp. 178-204). Cambridge: Cambridge University Press.

Childs, G. T. (2001). Research on ideophones, whither hence?: The need for a social theory of ideophones. In F. K. E. Voeltz & C. Kilian-Hatz (Eds.), *Ideophones* (pp. 63-73). Amsterdam/Philadelphia: John Benjamins.

Clift, R., & Holt, E. (2007). Introduction. In E. Holt & R. Clift (Eds.), *Reporting talk: Reported speech in interaction* (pp. 1-15). Cambridge: Cambridge University Press.

Couper-Kuhlen, E. (2012). Exploring affiliation in the reception of conversational complaint stories. In A. Peräkylä & M.-L. Sorjonen (Eds.), *Emotion in interaction* (pp. 113-146). Oxford: Oxford University Press.

Couper-Kuhlen, E., & Selting, M. (1996). *Prosody in conversation*. Cambridge: Cambridge University Press.

Crystal, D. (2008). *A dictionary of linguistics and phonetics* (6th Ed.). Malden, Oxford: Blackwell Publishing

Day, D., & Wagner, J. (2014). Objects as tools for talk. In M. Navile, P. Haddington, T. Heinemann, & M. Rauniomaa (Eds.), *Interacting with objects: Language, materiality, and social activity* (pp. 101-124). Amsterdam/Philadelphia: John Benjamins.

de Jong, N. (2001). The ideophone in Didnga. In F. K. E. Voeltz & C. Kilian-Hatz (Eds.), *Ideophones* (pp. 121-138). Amsterdam/Philadelphia: John Benjamins.

de Ruiter, J. P. (2000). The production of gesture and speech. In D. McNeill (Ed.), *Language and gesture* (pp. 284-311). Cambridge: Cambridge University Press.

de Ruiter, J. P. (2003). The function of hand gesture in spoken conversation In M. Bickenback, A. Klappert, & H. Pompe (Eds.), *Manus loquens* (pp. 338-347). Cologne: DuMont.

Dingemanse, M. (2011). *The meaning and use of ideophones in Siwu.* (Unpublished doctoral dissertation), Radboud University, Nijmegen.

Dingemanse, M. (2012). Advances in the cross-linguistic study of ideophones. *Language and Linguistics Compass, 6*(10), 654-672.

Dingemanse, M. (2013). Ideophones and gesture in everyday speech. *Gesture, 13*(2), 143-165.

Dingemanse, M. (2014). Making new ideophones in Siwu: Creative depiction in conversation. *Pragmatics and Society, 5*(3), 384–405.

Doke, C. M. (1935). *Bantu linguistic terminology*. London: Longmans.

Drew, P. (1998). Complaints about transgressions and misconduct. *Research on Language and Social Interaction, 31*(3&4), 295-325.

Drew, P., & Holt, E. (1988). Complainable matters: The use of idiomatic expressions in making complaints. *Social Problems, 35*(4), 398-417.

Ekström, A., & Lindwall, O. (2014). To follow the materials: The detection, diagnosis and correction of mistakes in craft education. In M. Navile, P. Haddington, T. Heinemann, & M. Rauniomaa (Eds.), *Interacting with objects: Language, materiality, and social activity* (pp. 227-248). Amsterdam/Philadelphia: John Benjamins.

Ford, C. E. (1993). *Grammar and interaction: Adverbial clauses in American English conversation*. Cambridge: Cambridge University Press.

Ford, C. E., & Fox, B. A. (2010). Multiple practices for constructing laughables. In D. Barth-Weingarten, E. Reber, & M. Selting (Eds.), *Prosody in interaction* (pp. 339-368). Amsterdam/Philadelphia: John Benjamins.

Ford, C. E., Fox, B. A., & Thompson, S. (Eds.). (2002). *The language of turn and sequence*. New York: Oxford University Press.

Ford, C. E., Thompson, S., & Drake, V. (2012). Bodily-visual practices and turn continuation. *Discourse Processes, 49*, 192-212.

Fortune, G. (1962). *Ideophones in Shona: An inaugural lecture given in the University College of Rhodesia and Nyasaland on 28 April 1961.* New York: Oxford University Press.

Fox, B. A., & Robles, J. (2010). It's like mmm: Enactments with it's like. *Discourse Studies, 12*(6), 715-738.

Fox, B. A., Thompson, S., Ford, C. E., & Couper-Kuhlen, E. (2012). Conversation analysis and linguistics. In J. Sidnell & T. Stivers (Eds.), *The handbook of Conversation Analysis* (pp. 726-739). Malden, MA: Wiley-Blackwell.

Glenn, P. (2003). *Laughter in interaction*. Cambridge: Cambridge University Press.

Goffman, E. (1978). Response cries. *Language, 54*(4), 787-815.

Goffman, E. (1981). *Forms of talk*. Philadelphia: University of Pennsylvania Press.

Goldin-Meadow, S., & Sandhofer, C. (1999). Gestures convey substantive information about a child's thoughts toordinary listeners. *Developmental Science, 2*, 67-74.

Goldin-Meadow, S., Singleton, J., & McNeill, D. (1996). Silence is liberating: Removing the handcuffs on grammatical expression in the manual modality. *Psychological Review, 103*(1), 34-55.

Goldin-Meadow, S., Wein, D., & Chang, C. (1992). Assessing knowledge through gesture: Using children's hands to read their minds. *Cognition and Instruction, 9*(3), 201-219.

Goodwin, C. (1979). The interactive construction of a sentence in natural conversation. In G. Pasthas (Ed.), *Everyday language: Studies in ethnomethodology* (pp. 97-121). N.Y.: Irvington.

Goodwin, C. (1981). *Conversational organization: Interaction between speakers and hearers*. NY: Academic Press.

Goodwin, C. (1984). Notes on story structure and the organiztaion of participation. In M. Atkinson & J. Heritage (Eds.), *Structures of social action* (pp. 225-246). Cambridge: Cambridge University Press.

Goodwin, C. (1986). Gestures as a resource for the organization of mutual orientation. *Semiotica, 62*(1/2), 29-49.

Goodwin, C. (2000). Action and embodiment within situated human interaction. *Journal of Pragmatics, 32*, 1489-1522.

Goodwin, C. (2007). Interactive footing. In E. Holt & R. Clift (Eds.), *Reporting talk: Reported speech in interaction* (pp. 16-46). Cambridge: Cambridge University Press.

Goodwin, C., & Goodwin, M. H. (1986). Gesture and coparticipation in the activity of searching a word. *Semiotica, 62*(1/2), 51-75.

Goodwin, M. H. (1983). Searching for a word as an interactive activity *Semiotics*, 129-138.

Goodwin, M. H., Cekaite, A., & Goodwin, C. (2012). Emotion as stance. In M.-L. Sorjonen & A. Peräkylä (Eds.), *Emotion in interaction* (pp. 16-41). Oxford: Oxford University Press.

Günther, S. (1997). Complaint stories: Constructing emotional reciprocity among women. In H. Kotthoff & R. Wodak (Eds.), *Communicating gender in context* (pp. 179-218). Amsterdam, NL: John Benjamins.

Haakana, M. (2007). Reported thought in complaint stories. In E. Holt & R. Clift (Eds.), *Reporting talk: Reported speech in interaction* (pp. 150-178). Cambridge: Cambridge University Press.

Haakana, M. (2010). Laughter and smiling: Notes on co-occurrences. *Journal of Pragmatics, 42*, 1499-1512.

Hamano, S. (1998). *The sound-symbolic system of Japanese*. Tokyo: Kurosio.

Hasada, R. (2001). Meanings of Japanese sound-symbolic emotion words. In J. Harkins & A. Wierzbicka (Eds.), *Emotions in crosslinguistic perspective* (pp. 217-253). Berlin, New York: Mouton de Gruyter.

Hayashi, M. (2003a). *Joint utterance construction in Japanese conversation*. Amsterdam/Philadelphia: John Benjamins.

Hayashi, M. (2003b). Language and the body as resources for collaborative action: A study of word searches in Japanese conversation. *Research on Language and Social Interaction, 36*(2), 109-141.

Hayashi, M. (2004). Discourse within a sentence: An exploration of postpositions in Japanese as an interactional resource. *Language in Society, 33*, 343-376.

Hayashi, M. (2005a). Joint turn construction through language and the body: Notes on embodiment in conjoined participation in situated activities. *Semiotica, 156*(1/4), 21-53.

Hayashi, M. (2005b). Referential problems and turn construction: An exploration of an intersection between grammar and interaction. *Text, 25*(4), 437-468.

Hayashi, M. (2009). Marking a 'noticing of departure' in talk: *Eh*-prefaced turns in Japanese conversation. *Journal of Pragmatics, 41*(10), 2100-2129.

Hayashi, M. (2012). Claiming uncertainty in recollection: A study of kke-marked utterances in Japanese Conversation. *Discourse Processes, 49*(5), 391-425.

Heritage, J. (1984). A change of state token and aspects of its sequential placement. In D. Atkinson & J. Heritage (Eds.), *Structures of social action* (pp. 299-345). Cambridge: Cambridge University Press.

Heritage, J., & Watson, R. (1979). Formulations as conversational objects. In G. Psathas (Ed.), *Everyday language: Studies in ethnomethodology* (pp. 123-162). New York, London: Irvington Press.

Hida, Y., & Asada, H. (2002). *Gendai giongo gitaigo yoohoo jiten* [Dictionary of modern use of phenomimies and phenomimes]. Tokyo: Tokyodo Shuppan.

Hinton, L., Nichols, J., & Ohala, J. J. (Eds.). (1994). *Sound symbolism*. Cambridge: Cambridge University Press.

Holt, E., & Clift, R. (Eds.). (2007). *Reporting talk: Reported speech in interaction*. Cambridge: Cambridge University Press.

Iverson, J. M., & Goldin-Meadow, S. (1998). Why people gesture as they speak. *Nature, 396*, 228.

Iwasaki, N., Brown, L., Kita, S., & Vinson, D. (2013). Japanese and Korean speakers' production of mimetic words. *Japanese/Korean Linguistics, 20*, 199-213.

Iwasaki, N., Vinson, D., & Vigliocco, G. (2007). What do English speakers know about *gera-gera* and *yota-yota*?: A cross-linguistic investigation of mimetic words for laughing and walking. *Sekai no nihongo kyouiku* [Japanese education around the world]*, 17*, 53-78.

Jefferson, G. (1985). An exercise in the transcription and analysis of laughter *Handbook of discourse analysis* (Vol. 3, pp. 25-34). London: Academic Press.

Jefferson, G. (1988). On the sequential organization of troubles talk in ordinary conversation. *Social Problems, 35*(4), 418-442.

Jefferson, G. (2004). Glossary of transcript symbols with an introduction. In G. Lerner (Ed.), *Conversation analysis: Studies from the first generation.* (pp. 13-31). Amsterdam/Philadelphia: John Benjamins.

Jefferson, G. (2015). *Talking about troubles in conversation* (P. Drew, J. Heritage, G. Lerner, & A. Pomerantz Eds.). Oxford: Oxford University Press.

Johnson, M. R. (1976). Toward a definition of the ideophone in Bantu. *Ohio State University Working Papers in Linguistics, 21*, 240-253.

Kakehi, H., & Tamori, I. (Eds.). (1993). *Onomatopia: Gion/gitaigo-no rakuen* [Onomatopoeia: A utopia of mimetics]. Tokyo: Keiso Shobo.

Kakehi, H., Tamori, I., & Schourup, L. (1996a). *Dictionary of iconic expressions in Japanese: A-J*. Berlin: Mouton de Gruyter.

Kakehi, H., Tamori, I., & Schourup, L. (1996b). *Dictionary of iconic expressions in Japanese: K-Z*. Berlin: Mouton de Gruyter.

Keevallik, L. (2014). Having a ball: Immaterial objects in dance instruction. In M. Navile, P. Haddington, T. Heinemann, & M. Rauniomaa (Eds.), *Interacting with objects: Language, materiality, and social activity* (pp. 249-268). Amsterdam/Philadelphia: John Benjamins.

Kendon, A. (1972). Some relationships between body motion and speech: An analysis of an example. In A. Siegman & B. Pope (Eds.), *Studies in dyadic communication* (pp. 177-210). Elmsford, NY: Pergamon.

Kendon, A. (1975). Gesticulation, speech, and the gesture theory of language. *Sign Language Studies, 9*, 349-373.

Kendon, A. (1980). Gesticulation and speech: Two aspects of the process of utterance. In M. R. Key (Ed.), *The relationship of verbal and nonverbal communication* (pp. 207-227). NY: Mouton de Gruyter.

Kendon, A. (1988). How gesture can become like words. In F. Poyatos (Ed.), *Cross-cultural perspective in nonverbal communication* (pp. 131-141). Tronto: Hogrefe.

Kendon, A. (1990). *Conducting interaction: Patterns of behavior in focused interaction*. Cambridge: Cambridge University Press.

Kendon, A. (2004). *Gesture: Visible actions as utterance*. Cambridge: Cambridge University Press.

Kita, S. (1993). *Language and thought interface: A study of spontaneous gestures and Japanese mimetics.* (Unpublished doctoral dissertation), University of Chicago, Chicago.

Kita, S. (1997). Two-dimensional semantic analysis of Japanese mimetics. *Linguistics, 35*, 379-415.

Kita, S. (2000). How representational gestures help speaking. In D. McNeill (Ed.), *Language and gesture* (pp. 162-185). Cambridge: Cambridge University Press.

Kita, S. (2001). Semantic schism and interpretive integration in Japanese sentences with a mimetic: A reply to Tsujimura. *Linguistics, 39*(2), 419-436.

Kita, S. (2002). *Jesucha: Kangaeru karada* [Gesture: thinking body]. Tokyo: Kaneko Shobo.

Kita, S. (2008). World-view of protolanguage speakers as inferred from semanctics of sound symbolic words: A case of Japanese mimetics. In N. Masataka (Ed.), *The origins of language: Unraveling evolutionary forces* (pp. 25-38). Berlin: Springer.

Klamer, M. (2001). Expressives and iconicity in the lexicon. In F. K. E. Voeltz & C. Kilian-Hatz (Eds.), *Ideophones* (pp. 165-181). Amsterdam/Philadelphia: John Benjamins.

Kotthoff, H., & Wodak, R. (Eds.). (1997). *Communicating gender in context*. Amsterdam, NL: John Benjamins.

Krauss, R. M., Morrel-Samuels, P., & Colasante, C. (1991). Do conversational hand gestures communicate? *Journal of Personality and Social Psychology, 61*(5), 743-754.

Kunene, D. P. (2001). Speaking the act: The ideophone as a linguistic revel. In F. K. E. Voeltz & C. Kilian-Hatz (Eds.), *Ideophones* (pp. 183-191). Amsterdam/Philadelphia: John Benjamins.

Kuno, S. (1973). *The structure of the Japanese language*. Cambridge, Massachusetts: Massachusetts Institute of Technology Press.

Labov, W. (1984). Intensity. In D. Schiffrin (Ed.), *Georgetown University Round Table on Languages and Linguistics* (pp. 43-70). Washington, D.C.: Georgetown University Press.

LeBaron, C. D., & Streeck, J. (2000). Gestrues, knowledge, and the world. In D. McNeill (Ed.), *Language and gesture* (pp. 118-138). Cambridge: Cambridge University Press.

Lerner, G. (1996). On the "semi-permeable" character of grammatical units in conversation: Conditional entry into the turn space of another speaker. In E. Ochs, E. A. Schegloff, & S. Thompson (Eds.), *Interaction and grammar*. Cambridge: Cambridge University PRess.

Lerner, G. (2002). Turn-sharing: The choral co-production of talk-in-interaction. In C. E. Ford, B. A. Fox, & S. Thompson (Eds.), *The language of turn and sequence* (pp. 225-256). Oxford: Oxford University Press.

Levinson, S. C. (2012). Action formation and ascription. In J. Sidnell & T. Stivers (Eds.), *The handbook of Conversation Analysis* (pp. 103-130). Malden, MA: Wiley-Blackwell.

Levy, E. T., & Fowler, C. A. (2000). The role of gestures and other graded language forms in the grounding of reference in perception. In D. McNeill (Ed.), *Langauge and gesture* (pp. 215-234). Cambridge: Cambridge University Press.

Lyons, J. (1977). *Semantics*. Cambridge: Cambridge University Press.

Mandelbaum, J. (1991/1992). Conversational non-cooperation: An exploration of disattended complaints. *Research on Language and Social Interaction, 25*, 97-138.

Masataka, N. (Ed.) (2008). *The origins of language: Unraveling evolutionary forces*. Berlin: Springer.

Matthews, B. (2014). Trajectories of the object in interaction. In M. Navile, P. Haddington, T. Heinemann, & M. Rauniomaa (Eds.), *Interacting with objects: Language, materiality, and social activity* (pp. 381-388). Amsterdam/Philadelphia: John Benjamins.

Mayberry, R. I., & Jaques, J. (2000). Gesture production during stuttered speech: Insights into teh nature of gesture-speech integration. In D. McNeill (Ed.), *Language and gesture* (pp. 199-214). Cambridge: Cambridge University Press.

McGreger, W. (2001). Ideophones as the source of verbs in Northern Australian languages. In F. K. E. Voeltz & C. Kilian-Hatz (Eds.), *Ideophones* (pp. 205-221). Amsterdam/Philadelphia: John Benjamins.

McNeill, D. (1992). *Hand and mind: What gestures reveal about thought*. Chicago: University of Chicago Press.

McNeill, D. (Ed.) (2000). *Language and gesture*. Cambridge: Cambridge University Press.

McNeill, D., Cassell, J., & McCullough, K. (1994). Communicative effects of speech-mismatched gestures. *Research on Language and Social Interaction, 27*, 223-237.

Mikkola, P., & Lehtinen, E. (2014). Initiating activity shifts through use of appraisal forms as material objects during performance appraisal interview. In M. Navile, P. Haddington, T. Heinemann, & M. Rauniomaa (Eds.), *Interacting with objects: Language, materiality, and social activity* (pp. 57-78). Amsterdam/Philadelphia: John Benjamins.

Miura, A., & McGloin, N. H. (2008). *An integrated approach to intermediate Japanese* (2nd ed.). Tokyo: The Japan Times.

Mondada, L. (2006). Video recording as the reflexive preservation and configuration of phenomenal features for analysis. In H. Knoblauch, J. Raab, S. H.-G., & B. Schnettler (Eds.), *Video analysis* (pp. 1-18). Bern: Lang.

Mondada, L. (2012). Talking and driving: Multiactivity in the car. *Semiotica, 2012*(191), 223-256.

Mondada, L. (2014). Cooking instructions and the shaping of things in the kitchen. In M. Navile, P. Haddington, T. Heinemann, & M. Rauniomaa (Eds.), *Interacting with objects: Language, materiality, and social activity* (pp. 199-226). Amsterdam/Philadelphia: John Benjamins.

Mori, J. (1999). *Negotiating agreement and disagreement in Japanese: Connective expressions and turn construction*. Amsterdam/Philadelphia: John Benjamins.

Mori, J. (2006). The workings of the Japanese token *hee* in informing sequences: An analysis of sequential context, turnshape, and prosody. *Journal of Pragmatics, 38*, 1175-1205.

Mori, J., & Hayashi, M. (2006). The achievement of intersubjectivity through embodied completions: A study of interactions between first and second language speakers. *Applied Linguistics, 27*(2), 195-219.

Morrel-Samuels, P., & Krauss, R. M. (1992). Word familiarity predicts temporal asynchrony of hand gestures and speech. *Journal of Experimental Psychology, 18*(3), 615-622.

Navile, M., Haddington, P., Heinemann, T., & Rauniomaa, M. (2014a). On the interactional ecology of objects. In M. Navile, P. Haddington, T. Heinemann, & M. Rauniomaa (Eds.), *Interacting with objects: Language, materiality, and social activity* (pp. 3-26). Amsterdam/Philadelphia: John Benjamins.

Navile, M., Haddington, P., Heinemann, T., & Rauniomaa, M. (Eds.). (2014b). *Interacting with objects: Language, materiality, and social activity*. Amsterdam/Philadelphia: John Benjamins.

Nielsen, S. b. (2014). "I'll just see what you had before": making computer use relevant while patients present their problems. In M. Navile, P. Haddington, T. Heinemann, & M. Rauniomaa (Eds.), *Interacting with objects: Language, materiality, and social activity* (pp. 79-97). Amsterdam/Philadelphia: John Benjamins.

Nobe, S. (2000). Where do most spontaneous representationalgestures actually occur with respect to speech? In D. McNeill (Ed.), *Language and gesture* (pp. 186-198). Cambridge: Cambridge University Press.

Noss, P. A. (2003). Ideophones. In P. M. Peek & K. Yankah (Eds.), *African folklore: An encyclopedia* (pp. 180-181). London: Routledge.

Ochs, E. A., Schegloff, E. A., & Thompson, S. A. (Eds.). (1996). *Interaction and grammar*. Cambridge: Cambridge University Press.

Ohala, J. J. (1994). The frequency code underlies the sound-symbolic use of voice pitch. In L. Hinton, J. Nichols, & J. J. Ohala (Eds.), *Sound symbolism* (pp. 325-347). Cambridge: Cambridge University Press.

Okpewho, I. (1992). *African oral literature: Backgrounds, character, and continuity*. Indiana: Indiana University Press.

Olson, D. R., & Torrance, N. (1983). Writing and criticizing texts. In B. Kroll & G. Wells (Eds.), *Explorations in the development of writing* (pp. 31-42). Malden; MA: Wiley.

Ono, M. (2007). *Nihongo Onomatope Jiten* [Dictionary of Japanese Onomatopoeias]. Tokyo: Shogakukan.

Ono, T., & Suzuki, R. (1992). Word order varriability in Japanese conversation: Motivations and grammaticalization. *Text, 12*, 429-445.

Pasthas, G. (Ed.) (1979). *Everyday language: Studies in ethnomethodology*. New York, London: Irvington Press.

Peräkylä, A. (2012). Epilogue: What does the study of interaction offer to emotion research? In A. Peräkylä & M.-L. Sorjonen (Eds.), *Emotion in interaction* (pp. 274-289). Oxford: Oxford University Press.

Peräkylä, A., & Ruusuvuori, J. (2006). Facial expression in an assessment. In H. Knoblauch, B. Schnettler, J. Raab, & H.-G. Soeffner (Eds.), *Video analysis: Methodology and methods* (pp. 127-142). Frankfurt: Peter Lang.

Peräkylä, A., & Ruusuvuori, J. (2012). Facial expression and interactional regulation of emotion. In A. Peräkylä & M.-L. Sorjonen (Eds.), *Emotion in interaction* (pp. 64-91). Oxford: Oxford University Press.

Peräkylä, A., & Sorjonen, M.-L. (Eds.). (2012). *Emotion in interaction*. Oxford: Oxford University Press.

Perlman, M. (2010). Talking fast: The use of speech rate as iconic gesture. In F. Perrill, V. Tobin, & M. Turner (Eds.), *Meaning, form and body* (pp. 245-262). Stanford: CSLI Publications.

Perrill, F., Tobin, V., & Turner, M. (Eds.). (2010). *Meaning, form and body*. Standord: CSLI Publications.

Prior, M. T. (2016). Formulating and scaling emotionality in L2 qualitative research interviews. In M. T. Prior & G. Kasper (Eds.), *Emotion in multilingual interaction* (pp. 203–236). Amsterdam/Philadelphia: John Benjamins.

Richardson, E., & Stokoe, E. (2014). The order of ordering: Objects, requests and embodied conduct in a public bar. In M. Navile, P. Haddington, T. Heinemann, & M. Rauniomaa (Eds.), *Interacting with objects: Language, materiality, and social activity* (pp. 31-56). Amsterdam/Philadelphia: John Benjamins.

Ruusuvuori, J., & Peräkylä, A. (2009). Facial and verbal expressions in assessing storis and topics. *Research on Language and Social Interaction, 42*(4), 377-394.

Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. *Language, 50*, 696-735.

Sakai, S., Korenaga, R., Mizukawa, Y., & Igarashi, M. (2014). Envisioning the plan in interaction: Configuring pipes during a plumbers' meeting. In M. Navile, P. Haddington,

T. Heinemann, & M. Rauniomaa (Eds.), *Interacting with objects: Language, materiality, and social activity* (pp. 339-356). Amsterdam/Philadelphia: John Benjamins.

Samarin, W. J. (1971). Survey of Bantu ideophonoes. *African Language Studies, 12*, 130-168.

Scheflen, A. E. (1963). Communication and regulation in psychotherapy. *Psychiatry, 26*(2), 126-136.

Scheflen, A. E. (1964). The significance of posture in communication systems. *Psychiatry, 27*(4), 316-331.

Schegloff, E. A. (1984). On some gestures' relation to talk. In M. Atkinson & J. Heritage (Eds.), *Structures of social action* (pp. 266-296). Cambridge: Cambridge University Press.

Schegloff, E. A. (1997). Practices and actions: Boundary cases of other-initiated repair. *Discourse Processes, 23*, 499-545.

Schegloff, E. A., & Sacks, H. (1973). Opening up closings. *Semiotica, 7*, 289-327.

Schourup, L. (1993). *Nihongo no kakikotoba hanashikotoba ni okeru onomatope no bunpu ni tsuite* [On distribution of onomatopoeia in written and spoken languages in Japanese]. In H. Kakehi & I. Tamori (Eds.), *Onomatopia* [Onomatopoeia] (pp. 77-100). Tokyo: Sokei Shobo.

Selting, M. (2010). Affectivity in conversational storytelling: An analysis of displays of anger or indignation in complaint stories *Pragmatics, 20*(2), 229-277.

Selting, M., & Couper-Kuhlen, E. (2001). *Studies in interactional linguistics*. Amsterdam/Philadelphia: John Benjamins.

Shibatani, M. (1990). *The languages of Japan*. Cambridge: Cambridge University Press.

Shimotani, M. (2008). *An analysis of the reactive token eh in Japanese conversation.* (Unpublished doctoral dissertation), University of Wisconsin-Madison, Madison, WI.

Shinohara, K., & Uno, R. (2013). *Onomatope kenkyuu no shatei: Chikazuku oto to imi* [Sound symbolism and mimetics: Rethinking the relationship between sound and meaning in language]. Tokyo: Hitsuji Shobo.

Shintel, H., Nusbaum, H. C., & Okrent, A. (2006). Analog acoustic expression in speech communication. *Journal of Memory and Language, 55*, 167-177.

Sidnell, J. (2006). Coordinating gesture, talk, and gaze in reenactments. *Research on Language and Social Interaction, 39*(4), 377-409.

Sidnell, J. (2010). *Conversation analysis: An introduction*. Malden, MA: Wiley-Blackwell.

Sidnell, J., & Stivers, T. (Eds.). (2012). *The handbook of Conversation Analysis*. Malden, MA: Wiley-Blackwell.

Sorjonen, M.-L. (2001). *Responing in conversation: A study of response particles in Finnish*. Amsterdam/Philadelphia: John Benjamins

Sorjonen, M.-L., & Peräkylä, A. (2012). Introduction. In A. Peräkylä & M.-L. Sorjonen (Eds.), *Emotion in interaction* (pp. 3-15). Oxford: Oxford University Press.

Streeck, J. (1993). Gestures as communication I: Its coordication with gaze and speech. *Communication monographs, 60*, 275-299.

Streeck, J. (1994). Gesture as communication II: The audience as co-author. *Research on Language and Social Interaction, 27*(3), 239-267.

Streeck, J. (2002). Grammars, words, and embodied meanings: On the uses and evlution of *so* and *like*. *Journal of Communication, 52*(3), 581-596.

Streeck, J., Goodwin, C., & LeBaron, C. D. (2011). *Embodied interaction: Language and body in the material world*. Cambridge: Cambridge University Press.

Tamori, I. (1993). *Nihongo onomatope no on'in keitai* [Phoneme morphology of Japanese onomatopoeia] In H. Kakehi & I. Tamori (Eds.), *Onomatopia: Gion/gitai-no rakuen* [Onomatopoeia: A utopia of mimetics] (pp. 1-15). Tokyo: Keiso Shobo.

Tamori, I. (2002). *Onomatope: Gion gitai go o tanoshimu* [Onomatopoeia: Appreciate phonomimes and phenomimes]. Tokyo: Iwanami Shoten.

Tamori, I., & Schourup, L. (1999). *Onomatope: Keitai to imi* [Onomatopoeia: Morphology and meaning]. Tokyo: Kuroshio.

Tanaka, H. (1999). *Turn-taking in Japanese conversation: A study in grammar and interaction*. Amsterdam/Philadelphia: John Benjamins.

Tanaka, H. (2000). Turn projection in Japanese talk-in-interaction. *Research on Language and Social Interaction, 33*(1), 1-38.

Tanaka, H. (2001). Adverbials for turn projection in Japanese: Toward a demystification of the "telepathic" mode of communication. *Language in Society, 30*, 559-587.

Toratani, K. (2005). A cognitive approach to mimetic aspect in Japanese. *Proceedings of the Annual Meeting of the Berkeley Linguistic Society, 31*(1), 335-346.

Toratani, K. (2006). On the optionality of to-marking on reduplicated mimetics in Japanese. *Japanese/Korean Linguistics, 14*, 415-422.

Toratani, K. (2009). Translating mimetics in Japanese: A cognitive approach *New Voices in Translation Studies, 5*, 63-77.

Voeltz, F. K. E., & Kilian-Hatz, C. (2001a). Introduction. In F. K. E. Voeltz & C. Kilian-Hatz (Eds.), *Ideophones* (pp. 1-8). Amsterdam/Philadelphia: John Benjamins.

Voeltz, F. K. E., & Kilian-Hatz, C. (Eds.). (2001b). *Ideophones*. Amsterdam/Philadelphia: John Benjamins.

Watson, R. L. (2001). A comparison of some southeast Asian ideophones with some African ideophones. In F. K. E. Voeltz & C. Kilian-Hatz (Eds.), *Ideophones* (pp. 385-406). Amsterdam/Philadelphia: John Benjamins.

Wilkinson, S., & Kitzinger, C. (2006). Surprise as an interactional achievement: Reaction tokens in conversation. *Social Psychology Quarterly, 69*(2), 150-182.

Wootton, A. J. (2012). Distress in adult–child interaction. In A. Peräkylä & M.-L. Sorjonen (Eds.), *Emotion in Interaction* (pp. 42-63). Oxford: Oxford University Press.