Context-Containing Constructions:
Why can we interpret an expression infinitely?

Taro Okahisa
Kyoto University
oka9taro@gmail.com

Keywords: context; construction; polysemy; multimodality

1. Introduction

The objective of this paper is to deal with what I call the infinite-list problem of meaning of a lexical item by introducing the notion of *context-containing construction*. Without the concept of polysemy, the device can explain why a single expression is interpreted as different meanings. This paper illustrates cognitive-psychological validity based on discussion in Taylor (2012) and Langacker (2001), and demonstrates broad applicability of context-containing construction through considering some case studies. Because context-containing construction is involved with Taylor's explanation of polysemy (cf. Taylor 2012: Ch.10), we will begin by briefly reviewing his idea in the next section. In Section 3, we will then define a context-containing construction. Finally, the notion's applicability to a broad variety of phenomena will be demonstrated in Section 4.

2. Polysemy and Contextual Profile

When considering meanings of a word, we need to reflect on the number of them. In the following examples, the word *open* appears to have three meanings at least.

(1) a. open the door
b. open your mouth
c. open the book

(Taylor 2012: 227)

The example (1a) expresses twisting the door knob and pushing or pulling it. In (1b), the activity involves only a part of the body, *your mouth*. The phrase (1c) signifies turning the cover of a book.

It is well-known that there are two approaches to explain the difference in meanings of a word as in (1): the lumping approach and the splitting approach (cf. Dirven 2005, Taylor 2012). The lumping approach assumes that a word has a single schematic meaning and that both semantic and pragmatic rules are applied to an utterance in order to specify the meaning of each word comprising it. By contrast, the splitting approach accepts "a many-to-one relation of meanings to forms" (Taylor 2012: 233). Although both approaches are fundamentally different, they have the same problem, that is, what I call the infinite-list problem (cf. Clark 1983, Pustejovsky 1995,

Allwood 2002, Taylor 2012). Whether we employ the lumping view or the splitting one, we need the infinite list of rules/meanings because a specific meaning of a word used in a sentence varies according to the context.

Taylor does not agree with both approaches and posits that "the units of knowledge are not the individual verbs *per se*, but the typical collocations of the verbs and the kinds of situations that they can refer to" (Taylor 2012: 227), and designates the environment around a word as "*the contextual profile*," as cited below:

> Words do not stand for so many discrete chunks of semantic content and speakers certainly do not learn the words of their language by learning these form-meaning associations. Rather, words need to be characterized in terms of their CONTEXTUAL PROFILE, having to do with the kinds of items they collocate with, the constructions they feature in, and the kinds of texts they occur in, in addition to their purely conceptual value and their referential possibilities.
>
> (Taylor 2012: 220)

Moreover, Taylor assumes two dimensions of contextualization of words and their meanings: the linguistic context and the conceptual context. The linguistic context is the collocation with a word, which is, as we have seen, essential to modulate the meaning of a word. The conceptual context is a set of cognitive domains which is activated in interpreting the word.[1] Consider the word *snail* as an example. *Snail* will remind us of rain, if we regard the conceptual context as the domain of garden pests. It will also invoke french restaurants, if the context is taken as the domain of fancy cuisine (cf. Langacker 2008: 49). This example shows that the meaning of a word is dependent on the conceptual context against which it is understood.

Taylor's contextual profile can describe polysemy of words, assuming that linguistic units include the linguistic and the conceptual context. His approach, however, is not able to completely explain polysemy of sentences in discourse. Consider, as an example, the ambiguous sentence in (2).

(2)

Mom: Do you know who's going to that meeting?
Kid: Who
Mom: I don't know!
Kid: Ouh:: prob'ly: Mrs. Murphy an' Dad said prob'ly Mrs. Timpte en some a' the teachers.

(Terasaki 2004 [1976]: 202)

In the example, the mother's first utterance was misunderstood by the kid because he took it as "pretelling" or "preannouncement" while she actually wanted to ask him about the participants of *that meeting*. This misunderstanding can be attributed to the difference in conceptual context on which the kid and his mother focus respectively, because the utterance was made in the same linguistic context. Without extra pragmatic rules, however, Taylor's approach cannot account for

why their conceptual contexts were different in spite of the fact that the they shared the same linguistic context in the same situation. In the following sections, we will introduce the notion of context-containing construction, which is able to explain this sort of misunderstanding completely from the perspective of construction theory.

3. The Definition of Context-Containing Constructions

The notion of context-containing construction is based on Langacker's "*symbolization*", which is defined as "the association of conceptualization with the mental representation of observable entities" (Langacker 2000: 5).[2] It should be noted that input of symbolization is mental imagery (i.e. mental representation) rather than a stimulus (i.e. observable entity) itself. Considering that we have a working memory, in which previous conceptualizations are stored, a preceding output can be a following input.[3] This process is depicted roughly in Figure 1. Stimuli (S) from the external world (EW) are converted into corresponding mental representations (R) in what I call mental desktop (MD), in which there are active or semi-active information (cf. Chafe 1994). The conversion is indicated by solid arrows. Some mental representations invoke a particular conceptualization (C), as represented by dashed arrows. The conceptualization can be analyzed into a set of mental representations activated by inputs. The representations of stimuli and the conceptualization are loaded into the working memory (WM) to interpret the following stimuli. At the time t_1, C_1 is the output associated with the set of R_{1a}, R_{1b} and R_{1c}. However, at the time t_2,

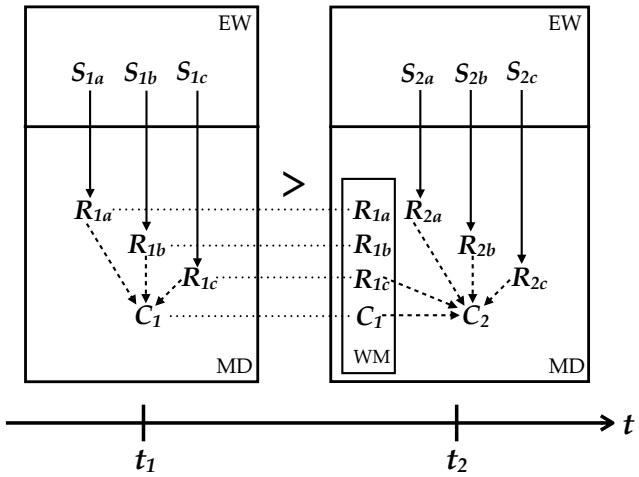

Figure.1 A flow of processing stimuli

it plays the role of an input for C_2. In this case, R_{1c} restored in the working memory also supports invoking C_2. The dotted line indicates the correspondence between t_1 and t_2.

Given that preceding outputs can play a role of following inputs, we can classify inputs of mental representations into three types in accordance with their characteristics:

(3) a. Phonemic Sequence (PS)
Mental representations of speech

b. Multimodal Information (MI)
Mental representations of physical stimuli, except for speech
(e.g. facial expression, gesture, sound, scent)

c. Accesible Information (AI)
Mental representations which are activated and accessible at a moment
(e.g. previous linguistic context, background knowledge)

Phonemic Sequence and Multimodal Information constitute "the mental representation of observable entities." Because linguistic elements are strongly associated with a particular representation, phonemic sequences are distinguished from multimodal information.[4] Accessible Information comprises previous inputs (R_{1a}, R_{1b} and R_{1c} in Figure 1) and what is activated by them (C_1 in Figure 1).[5] It is stored in the working memory, supporting activation of the following conceptualization. Some pieces of accessible information are rendered inaccessible according to one's attentional focus or to the limitations of working memory. MI and AI are regarded as context, which is psychological and linguistic environment on which one focus in order to interpret something.

Based on the above discussion, let us define context-containing constructions:

(4) Context-Containing Construction
Symbolic units composed of an input set of mental representations and an output set of those it activates

Following the definition, in the case of Figure 1, two context-containing constructions are found as follows:

(5) a. $\left[\{\text{PS/MI: } R_{1a}, R_{1b}, R_{1c}\} \rightarrow \{C_1\}\right]$

b. $\left[\left\{\begin{array}{c}\text{PS/MI: } R_{2a}, R_{2b}, R_{2c} \\ \text{AI: } R_{1c}, C_1\end{array}\right\} \rightarrow \{C_2\}\right]$

In the above examples, the right arrows indicate that the left-hand sides (i.e. input sets) invoke the right-hand sides (i.e. output sets). In the former construction (5a), which occurs at t_1, the input set comprises R_{1a}, R_{1b} and R_{1c}, invoking the output set, which is composed of C_1.[6] The elements of the input set are Accessible Information or Multimodal Information since they are derived from the observable entities at t_1. The latter construction (5b) is evoked at t_2, and means that the input set, whose members are R_{1c}, C_1, R_{2a}, R_{2b} and R_{2c}, activates the output set, that is, C_2. It should be

noted that R_{1c} and C_1 are Accessible Information in the working memory at t_2. In the following, I use this notation in order to describe context-containing constructions.

4. Phenomena of Context-Containing Constructions

If contexts (i.e. MI and AI) are added to the notion of construction, it can logically be assumed that there are seven types of context-containing construction according to combinations of inputs:

(6) a. PS construction (e.g. idiophone)
b. MI construction (e.g. phonogram)
c. AI construction (e.g. adjacent pair)
d. PS-MI construction (e.g. multimodal construction)
e. PS-AI construction (e.g. word, poetic rhetoric)
f. MI-AI construction (e.g. gesticulation)
g. PS-MI-AI construction (e.g. syntax-representing gesture)

In linguistic phenomena, constructions composed of Phonemic Sequence and Accessible Information in (6e) are attested frequently, because most of the linguistic elements are dependent on the context. It may therefore be useful to begin by considering PS-AI constructions in detail.

4.1 PS-AI Constructions

For the case of *snail* we have seen in Section 2, Taylor (2012) claims that which meaning is invoked depends on the conceptual context, the cognitive domain against which the word is interpreted. This explanation is of course appropriate, but it does not give an account for why a particular domain is activated. Context-containing constructions can simply explain by observing that the inputs of AI modulate their cognitive domains; the *snail* invoking rain and the *snail* invoking a french restaurant are two different constructions as described in (7).

(7) a. Snail-pest construction

$$\left[\left\{\begin{array}{l}\text{PS: /sneɪl/}\\ \text{AI: `the outdoor environment'}\end{array}\right\} \rightarrow \{\text{`snail as a kind of garden pest'}\}\right]$$

b. Snail-cuisine construction

$$\left[\left\{\begin{array}{l}\text{PS: /sneɪl/}\\ \text{AI: `food'}\end{array}\right\} \rightarrow \{\text{`snail as a kind of fancy cuisine'}\}\right]$$

If you have been talking about something in 'the outdoor environment,' you will employ the former construction (7a). The latter construction (7b) is, on the other hand, used when 'food' is activated at the time. The same goes for all content words because we could not specify the meaning of a word without the context around it.[7] Next, let us consider an example of the phrase,

eat snails. At first sight the word *eat* may appear to activate 'FOOD.' However, if the subject is *a dog*, what is relevant is the construction of (7a), because the AI in the construction of (7b) can be corrected to 'HUMAN FOOD' strictly speaking and the word *dog* activates the representation of 'THE OUTDOOR ENVIRONMENT.'

Basically, the whole sentence modulates the meaning of each word and phrase, playing a role of a trigger of a specific construction. In some cases, however, hearers cannot specify the construction which the speaker intends in spite of hearing the whole sentence. Recall the misunderstanding in (2). Although the kids and his mother recognized the same utterance, they took it as different meanings. It seems that they employ the following constructions.

(8) a. Preannouncement construction

$$\left[\begin{Bmatrix}\text{PS: /dʊ ju: noʊ ..... ?/}\\ \text{AI: 'AT THE BEGINNING OF A TOPIC' \& 'THE SPEAKER KNOWS .....'}\end{Bmatrix}\right.$$
$$\left.\rightarrow \{\text{'PREANNOUNCEMENT'}\}\right]$$

b. Request-for-information construction

$$\left[\begin{Bmatrix}\text{PS: /dʊ ju: noʊ ..... ?/}\\ \text{AI: 'THE SPEAKER DOES NOT KNOW ABOUT .....'}\end{Bmatrix}\right.$$
$$\left.\rightarrow \{\text{'REQUEST FOR INFORMATION'}\}\right]$$

In this case, the kid and his mother recognize the utterance in question to be the 'BEGINNING OF THE TOPIC.'[8] He seems to have misunderstood her facial expression or speech characteristics as meaning that 'SHE KNOWS WHO IS GOING TO THAT MEETING.'[9] He employed, therefore, the construction of (8a) in order to understand the mother's utterance although she actually used that of (8b). Thus, this misunderstanding can be attributed to the difference in the employed constructions between them. Yamanashi (2009) points out a kind of construction including contexts as the "*global construction*:"

(9) A: Drinking too much is not good for your health.
B: I wasn't born yesterday.
(Yamanashi 2009: 234)

In the case of (9), the utterance *I wasn't born yesterday* conveys the meaning of 'COMPLAINT' while it cannot do so when used by itself. Yamanashi defines global constructions as constructions whose meaning is dependent on previous discourse/text. He then argues that we learn these constructions as a interactive pattern rather than individual words or constructions.

Constructions composed of PS and AI are related to grammatical phenomena such as genre-based argument omissions. In some genres, an argument of a verb is omitted as in the following examples.

(10) a. ((In a recipe))
Sweet Lassi Ingredients:

1 Serving Plain yogurt
1 cup Sugar
2 tablespoons Ice Cubes
Method: Blend all the ingredients in an electric blender. Serve Ø cold.

b. ((In a product label))
Ø Contains alcohol.

c. ((In a report of a soccer game))
He hammered Ø wide of Gary Walsh's exposed net.

(Ruppenhofer and Michaelis 2010: 160)

Ruppenhofer and Michaelis (2010) offers a constructional account of this phenomenon. They claim that "[j]ust as reading, drinking and smoking conventionally involve certain types of participants, so recipes, product labels and match reports invoke a specific set of conventional participants (an edible substance, a product, a ball, etc.)" (Ruppenhofer and Michaelis 2010: 165). Their analysis is compatible with the notion of context-containing construction, because genre can be regarded as AI.

As a last example of PS-AI constructions, let us consider the rhetorical phenomena in poetics such as verse or paronomasia. There is a limited amount of work which illuminates poetic rhetoric with the aim of connecting it with other linguistic phenomena. Roman Jakobson is one of the few linguists who emphasize the continuum from linguistics to poetics. He argued that "the analysis of verse is entirely within the competence of poetics, and the latter may be defined as that part of linguistics which treats the poetic function in its relationship to the other functions of language" (Jakobson 1960: 359), defining "*the poetic function*" of language as " [t]he set (*Einstellung*) toward the message as such, focus on the message for its own sake" (Jakobson 1960: 356). From the point of view of paronomasia, he found the poetic characteristics in familiar phrases. For example, two words in the phrase ***in****nocent* ***by****stander* are dactylic like a waltz rhythm; and the famous Caesar's message *veni, vidi, vici* has "[t]he symmetry of three disyllabic verbs with an identical initial consonant and identical final vowel" (Jakobson 1960: 358). Employing context-containing construction, we can describe these rhetoric as follows:

(11) Dactyl construction

$$\left[\left\{\begin{array}{l}\text{PS: } /\acute{\sigma}\ \sigma\ \sigma\ /\\ \text{AI: } /\acute{\sigma}\ \sigma\ \sigma\ /\times n\end{array}\right\} \rightarrow \{\text{`DACTYL'}\}\right] (n \geqq 1)$$

(12) Alliteration construction

$$\left[\left\{\begin{array}{l}\text{PS: /O.../}\\ \text{AI: /O.../}\times n\end{array}\right\} \rightarrow \{\text{`ALLITERATION'}\}\right] (n \geq 1)$$

(13) Rhyme construction

$$\left[\left\{\begin{array}{l}\text{PS: /...R/}\\ \text{AI: /...R/} \times n\end{array}\right\} \rightarrow \{\text{`RHYME'}\}\right] (n \geqq 1)$$

In (11), the greek letter 'σ' represents any single syllable. It can be assumed, in my model, that AI includes not only previous outputs, but also previous inputs. The dactyl construction requires strong-weak-weak accent patterns as AI and the same pattern as PS. The capitals O and R in (12,13) indicate a specific phoneme. Both alliteration and rhyme constructions are composed of the PS identified with their AI. The meanings of these constructions are so abstract that we could not describe them without their name. However, the fact that we often recognize some utterances to be dactyl, alliteration or rhyme shows that we have such constructions.

What utterances do we feel to be poetic-rhetorical? For example, it is difficult to regard the following utterance as a poetic-rhetorical expression.

(14) Walker, runner, rider, and driver.

Although the utterance (14) corresponds with the rhyme construction (13), we do not take it as rhyme. This is because we also have the following construction and interpret the above utterance to be the result of repetition of it.

(15) Deverbal noun *-er* construction

$$\left[\left\{\begin{array}{l}\text{PS: /...ɚ/}\\ \text{AI: `\{/ ... /\} } \rightarrow \text{ \{`VERB'\}'}\end{array}\right\} \rightarrow \{\text{`AGENT THAT ...S'}\}\right]$$

Therefore, we recognize each word in (14) as an instance of the construction (15), which is, in Jakobson (1960), called "*equivalence*" on which the selection of words is based. Jakobson characterized the poetic function as projection of equivalence from selection into combination:

> *The poetic function projects the principle of equivalence from the axis of selection into the axis of combination*. Equivalence is promoted to the constitutive device of the sequence. In poetry one syllable is equalized with any other syllable of the same sequence; word stress is assumed to equal word stress, as unstress equals unstress; prosodic long is matched with long, and short with short; word boundary equals word boundary, no boundary equals no boundary; syntactic pause equals syntactic pause, no pause equals no pause. Syllables are converted into units of measure, and so are moras or stresses.
>
> (Jakobson 1960: 358)

From the perspective of context-containing construction, Jakobson's idea can be interpreted as following: poetic-rhetorical constructions require nothing but repetition of the same form. We could not feel an utterance to be poetic-rhetorical, when we recognize repetition of the same meaning. This characteristic of rhetorical constructions can be attributed to their AI elements. AI of rhetorical constructions is the previous input, while that of other constructions is the previous output. We would tend to focus on previous outputs more than previous inputs.

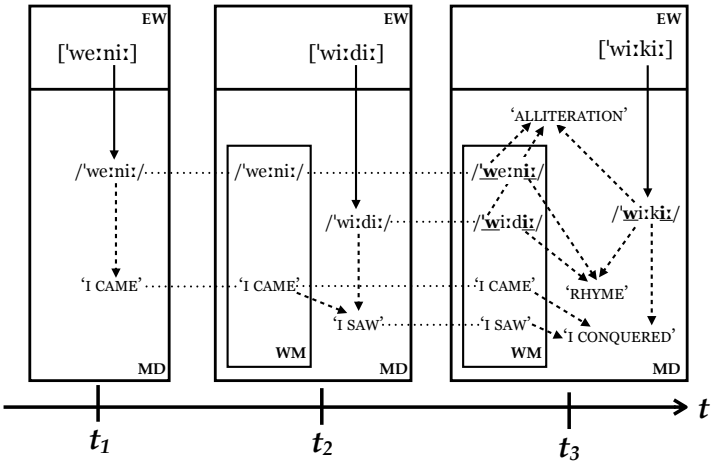

Figure.2 The procedure of interpreting *veni, veid, vici*

As an example of poetic-rhetorical constructions, let me consider the message *veni, veid, vici*. The procedure of invoking the constructions is abstractly shown in Figure 2.[10] At t_1, the PS /ˈwe:ni:/ invokes the conceptualization 'I CAME'. At the next moment t_2, they are stored in the working memory and play a role of AI. The PS /ˈwi:di:/ and the AI produce the conceptualization 'I SAW', which is the coherent interpretation that 'I' at t_2 is identified with 'I' at t_1. Finally, four mental representations are stored in the working memory as AI at t_3. Two of them are the previous inputs and the others are the previous outputs. Employing the former two representations (i.e. /ˈwe:ni:/ and /ˈwi:di:/) and the immediate input /ˈwi:ki:/, we can clearly recognize the existence of alliteration and rhyme in the message. By contrast, we can understand the whole meaning of the message by making use of the latter two representations (i.e. 'I CAME' and 'I SAW') and the immediate input.[11]

As we have seen above, there are many linguistic phenomena involving PS-AI constructions. This fact shows that contexts are essential for human language. In our communication system, however, there are some elements associated with conceptualization regardless of contexts. We will consider such constructions in the following two subsections.

4.2 PS Constructions

While we usually employ constructions containing contexts, I suspect that the usage of constructions without MI and AI are extremely limited. This is because there are few linguistic units

whose meaning is fixed and independent of contexts. One might suppose that proper nouns and grammatical elements can be regarded as examples of PS constructions. A moment's reflection, however, shows that they are also dependent on the context. Consider the following examples.

(16) a. Although I have not read his books, I love Haruki Murakami.
b. Although I do not like his characters, I love Haruki Murakami.

The proper noun *Haruki Murakami* in (16a) refers to the novelist. Conversely, *Haruki Murakami* in (16b) means books that Haruki Murakami wrote. This usage of proper nouns has been called metonymy in linguistics. These examples show that context is an essential factor even for the interpretation of proper nouns.

As an example of grammatical elements, let us consider the English plural suffix *-s*. At first sight, *-s* appears to mean 'PLURAL' regardless of contexts. However, the phonemic sequence (i.e. /s/, /z/ or /iz/) itself does not have any meaning. Do you interpret as a plural suffix all /z/ sounds in such an expression as *goes* and *Lyons*? These contradictions suggest that the construction of plural suffix should be described as the following:

(17) Plural construction

$$\left[\left\{\begin{array}{l}\text{PS: / ... \{s/z/iz\} /}\\ \text{AI: '\{/ ... /\} } \rightarrow \text{ \{'NOUN'\}'}\end{array}\right\} \rightarrow \{\text{'PLURAL'}\}\right]$$

This is a kind of PS-AI construction, which we saw in the previous subsection. Because most of grammatical elements have a short phonemic sequence, we could not identify them without context.

Then, are there no PS constructions in our communication system? In my view, conventional idiophones or mimetics can be regarded as a kind of PS construction. For English native speakers, *bow-wow* represents a dog's barking and *cock-a-doodle-doo* means a rooster's crow. By contrast, for Japanese speakers, *wanwan* corresponds to the former mimetic and *kokekokkoh* to the latter. We can, of course, verbalize the barking and the crow in another way, but such an imitation is difficult to understand without context. On the other hand, what we regard as idiophone invokes a particular meaning in any context because they have been conventionalized in each specific culture. In this point, PS constructions are similar to MI constructions, which we will consider in the next subsection.

4.3 MI Constructions

Just as PS constructions are agreed to be specific symbols in a culture, so MI constructions have been conventionalized in a particular community. A striking example of MI constructions is the traffic signal. In most of the developed countries, the green signal means 'SAFE' and the red 'UNSAFE'.[12] In addition, sirens of fire engines, police cars and ambulances can be categorized as a kind of MI construction because they are phonetic sounds rather than phonemes.[13]

As an example of MI constructions, consider the case of "*emblems*." Kendon explains that "they are used singly and as such they are the functional equivalent of a complete speech act" (Kendon 1988: 136). The circle made by a thumb and an index finger means 'OK' for American people regardless of context, for instance. It is worth noting that the gesture is, strictly speaking, not equal to the utterance '*okay*.' This is because we sometimes cannot perform the OK gesture when we can nevertheless say *okay* and vise versa. For example, the OK gesture can be impolite in some official meetings while we are allowed to say *okay* there. By contrast, if we are the on-site management staff in the meetings, we should use the OK gesture during the meeting in order to communicate between the staff members without voice. In this sense, the gesture has a particular meaning in any context.

Following the definition in Section 3, phonograms can also be regarded as MI constructions. For example, the Japanese phonogram (i.e. *hiragana*) 'あ' can be described as the following construction:

(18) Japanese *hiragana* /a/ construction

$$\left[\{\text{MI: [あ] }\} \rightarrow \{/\text{a}/\}\right]$$

Notice that the construction's output is a phoneme. In general, the output plays a role of a PS input in the following process of interpretation. Phonograms are similar to emblems in the sense that both are standardized in form, but we can express different nuances with different shapes. If we perform the OK gesture quickly, without hesitation, the connotation will be 'perfectly OK.' In the same way, the bold font [**あ**] means a loud voice or a stressed mora in some situations.[14]

MI constructions are attested in non-human languages as well. Some birds' singing to alert that an enemy is approaching can be regarded as MI constructions (perhaps PS constructions for the birds), which is a kind of custom shared by the members of a community.

4.4 AI Constructions

The context-containing constructions composed of AI can be triggered by multiple types of PS and MI. When people say to each other either '*hello*' or '*good morning,*' we can understand that they are meeting for the first time in that day. Each PS of both utterances invokes the common representation of greeting at the first encounter in a day, whereas they also activate different ones, which create the difference in meaning between them.[15]

We can find examples of such constructions in studies on discourse patterns. A striking example is, first of all, Propp's (1968) "*Morphology of the Folktale*." Propp was one of the most influential figures in the development of narratology. He listed 31 functions of folktales' dramatis personae by analyzing 100 tales. He found that a tale is composed of some functions in the list and they always occur in the order listed. Although these functions are not connected with any specific PS and MI, components of the list can be inputs invoking a plot including them. For example, the function XII (i.e. the hero is tested, interrogated, attacked, etc., which prepares the way for

his receiving either a magical agent or helper) reminds us of another function XIV (i.e. the hero acquires the use of a magical agent (e.g. animals; objects out of which magical helpers appear; objects possessing a magical property; qualities or capacities which are directly given)) and the whole particular story such as a Japanese folktale *Momotaro*, where the hero takes animals into his circle in order to fight marauding demons and finally defeating them. Such a domino effect shows that we have abstract patterns of stories as context-containing constructions.

The same goes for our daily conversation. A well-known example of a pattern of conversation is "*adjacency pair*" (Schegloff and Sacks 1973), which are basic pairs of utterances composing interactions such as 'question-answer,' 'greeting-greeting' and 'offer-acceptance/refusal.' Notice that, such as Propp's functions, an adjacency pair can be realized with various utterances.

These discourse patterns have mainly been considered in literary studies and conversation analysis. However, the notion of context-containing construction is able to integrate these studies with linguistic research.

4.5 PS-MI Constructions

In this subsection, we turn to constructions combining two types of input. Most of PS-MI constructions are identical with so called "*multimodal constructions*" (cf. Steen and Turner 2013). Studies on multimodal constructions demonstrate that "[t]he different modalities of a communicative act are not simply redundant, but provide new aspects of meaning" (Steen and Turner 2013: 19), based on research employing the Red Hen corpus, which is composed of audiovisual TV news broadcasts. For example, Zima (2014) points out that constructions invoking literal meaning co-occur with gestures more frequently than those invoking metaphorical meaning. She reports the frequency of the phrase *V(motion) in circles* accompanied by circular gestures as an example. According to her, the phrase employed as literal meaning co-occurrs with the gestures in 69% of cases while 37% of metaphorical uses are accompanied by the gestures. This study shows that constructions with the same PS can have different meanings depending on their MI.

It is worth noting that PS-MI constructions are difficult to distinguish from PS-AI constructions. This is because MI at a moment is stored in the working memory and then becomes AI in the following moment.[16] For convenience of discussion, however, we distinguish MI and AI depending on whether the information in question is a mental representation of an observable entity or not.

4.6 MI-AI Constructions

The boundary between MI-AI constructions and PS-MI constructions is not easy to draw since phonemic sequences composing a construction require time duration. In this paper, the former is defined as a pair of MI and a particular image, and the latter as one of MI and a particular PS. Following these definitions, '*gesticulation*,' which is invented by Kendon (1988), can be regarded

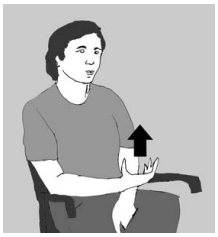

Figure.3 Gesture combining upward movement and interiority (McNeill 2005: 23)

as an example of MI-AI construction. Gesticulation can be defined as gestures integrated with speech ad hoc. McNeill (2000, 2005) characterizes it as the obligatory presence of speech and the absence of linguistic properties. While the former is attributed to the definition of gesticulation, the latter indicates that gesticulation does not involve phonological, morphemic and syntactic rules. In Figure 3, to take up one of McNeill's (2005: 23) examples, a woman uttering '*and he goes up thróugh the pipe this time*' is illustrated. She raises her hand upward co-expressively with '*up*' and makes her fingers spread outward to create an interior space with uttering '*through*.' In the gesture, she expresses the whole idea she wants to whereas she has to utter the two words, '*up*' and '*thróugh*' in the speech to convey the idea. McNeill adds that gesture and speech are not redundant because they express the underlying idea in distinctive ways. If her gesture in Figure 3 is performed without any context, we will not understand the meaning. Speech with gesticulation invokes, therefore, a different meaning from what is conveyed by the speech only or the gesture only. This example shows that we may have abstract constructions composed of MI and AI as in (6f).

(19) Gesticulation construction

$$\left[\left\{\begin{array}{l}\text{MI: [GESTURE]}\\ \text{AI: `WHAT [GESTURE] DESCRIBES'}\end{array}\right\} \rightarrow \{\text{`IMAGE AS A RESULT OF FUSING MI AND AI'}\}\right]$$

The gesticulation construction means that one can understand an immediate gesture as a specific meaning when linguistic context is compatible with what the gesture express. This perspective provides the following questions about gesticulation:

(20) a. What type of GESTURE is used in the construction?
b. What type of AI is described by GESTURE in the construction?
c. How long of a sequence can we recognize as a single construction?

Although the above questions have been addressed by studies on communication, they can be dealt with from the perspective of construction. The notion of context-containing construction provides a unified approach to explain both grammatical and communicational elements.

4.7 PS-MI-AI Constructions

Finally, let us consider gesture-disambiguation constructions as examples of PS-MI-AI constructions (Okahisa, in press). As we discussed above, grammatical interpretation is dependent on AI as well as on PS. We sometimes make use of MI in order to disambiguate utterances which can be interpreted in two ways. Okahisa (in press) demonstrates that speakers employ hierarchy-oriented and linearity-oriented gestures to disambiguate their utterances. In the experiment conducted there, the participants were shown two pictures and a Japanese ambiguous expression such as (21) on the computer screen (Figure 4).

(21) *Shikakui ryukku-no poketto* (四角い リュックの ポケット)
square backpack-GEN pocket
'[[The square backpack] pocket]' / '[The square [backpack pocket]]'

In front of the video camera, they were asked to utter the expression with gestures to convey the meaning which was emphasized by the bold frame around one of the two images. As indicated in Table 1, several gesture types significantly co-occurred with a particular meaning more frequently than the other meaning: the N1, N2, N1+N2 and N2+N1+N2 gesture types.[17]

The N1, N2 and N1+N2 types tend to be employed to represent grammatical heads in each hierarchy, as indicated in Figure 5. Since the structure [[Adj. N1] N2] has two heads, N1 and N2, the N1 and N1+N2 gesture types were performed when the utterance meant the structure. However, the utterances categorized as the N2 gesture type were observed only when the speakers tried to convey the structure [Adj. [N1 N2]], which has the only head N2. These gesture types can be said to be oriented to the hierarchy of sentences. On the other hand, the N2+N1+N2 type can be regarded as the representation of the connection between the first and the third word (i.e. Adj. and N2), considering that the first gesture corresponds with the third gesture. This type appears to focus on the linearity of sentences. These gesture types can be described as the following constructions:

Figure.4 An example of stimuli presented to the participants in Okahisa (in press)

Table.1 Frequencies in each gesture type (Okahisa, to appear)

	Gesture type						
Meaning	N1‡	N2‡	N1+N2†	N2+N1+N2†	Adj.+N1+N2	Adj.+N2	Adj.
[[Adj. N1] N2]	**23**	0	**25**	0	2	0	0
[Adj. [N1 N2]]	0	**29**	11	**7**	0	2	1

(The two meanings' frequencies in each gesture type were compared by the binomial test; † < .05,‡ <.001)

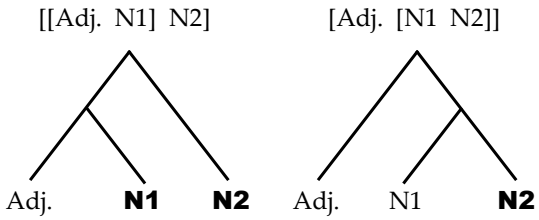

Figure.5 Hierarchal structures in [Adj. N1 N2]

(22) Gesture-hierarchy construction

a. $\left[\left\{\begin{array}{l}\text{PS: /...i X no Y/}\\ \text{MI: } [G_{n1}] \,/\, [G_{n1} + G_{n2}]\\ \text{AI: `\{/...i/\} → \{`ADJ'\}' \& `\{/X/\} → \{`N}_1\text{'\}' \& `\{/Y/\} → \{`N}_2\text{'\}'}\end{array}\right\} \rightarrow \{\text{`[[ADJ N}_1\text{] N}_2\text{]'}\}\right]$

b. $\left[\left\{\begin{array}{l}\text{PS: /...i X no Y/}\\ \text{MI: } [G_{n2}]\\ \text{AI: `\{/...i/\} → \{`ADJ'\}' \& `\{/X/\} → \{`N}_1\text{'\}' \& `\{/Y/\} → \{`N}_2\text{'\}'}\end{array}\right\} \rightarrow \{\text{`[ADJ [N}_1\text{ N}_2\text{]]'}\}\right]$

(23) Gesture-linearity construction

$$\left[\left\{\begin{array}{l}\text{PS: /...i X no Y/}\\ \text{MI: } [G_{n2} + G_{n1} + G_{n2}]\\ \text{AI: `\{/...i/\} → \{`ADJ'\}' \& `\{/X/\} → \{`N}_1\text{'\}' \& `\{/Y/\} → \{`N}_2\text{'\}'}\end{array}\right\} \rightarrow \{\text{`[ADJ [N}_1\text{ N}_2\text{]]'}\}\right]$$

In these constructions, N_1 and N_2 represent different nouns; G_{n1} and G_{n2} indicate gestures representing N1 and N2 respectively. It should be noted that the experiment forced the participants to perform gestures with uttering. In daily conversations, we might not use such gestures to disambiguate our utterances. There is, hence, a possibility that the result of the experiment is attributed to its condition. However, the fact that some constructions are frequently used in the experiment

means that they are motivated by cognitive factors. Context-containing construction can be said to be based not only on daily usage but also on our cognitive ability.

5. Conclusion

As we have seen, the notion of context-containing construction can explain various phenomena involving our interpretation, which has been examined in different disciplines: semantics, pragmatics, poetics, grammatology, narratology, multimodal analysis and so on. The overall picture of the phenomena considered in this paper is depicted in Figure 6. Note that it is not complete and there are many phenomena which are not handled here. However, it shows that diverse phenomena can be located in the continuum of context-containing construction.

Context-containing construction is, as mentioned previously, a device to avoid the infinite-list problem. One might object that the notion also requires an infinite list of constructions in order to explain various phenomena. There are, of course, many kinds of constructions involving interpretation of symbolic units. However, the constructions can be analyzed into basic and simple representations. Our goal is to unfold each composite construction and to reveal the mechanism behind its production and comprehension. For achievement of the goal, we have to investigate various phenomena from the same perspective and integrate the results.

Acknowledgments

This paper is based on a presentation at the 8th Conference on Language, Discourse and Cognition held at National Taiwan University in May 2016. I gratefully acknowledge the comments received there, and also from the reviewers of this paper. In addition, I would like to give special thanks to Jessica Tynes for her detailed comments on a preliminary version of this paper.

Notes

[1.] Taylor follows Langacker's usage of *cognitive domains*. Langacker (1987: 44) defines a cognitive domain as "[a] coherent area of conceptualization relative to which semantic units may be characterized."

[2.] Although my model is based on the "*symbolic unit*," which Langacker propose in his Cognitive Grammar, I use the term *construction* in this paper. This is because my model assumes that a single unit contains various pieces of information as an input. The term *construction* is employed to express multiple elements of the input.

[3.] In another paper, Langacker (2001) demonstrates that there are some symbolic units which are "*retrospective*, in the sense of making a specification concerning the prior discourse, and/or *prospective*, by virtue of evoking the subsequent discourse." He took up the Japanese topic marker

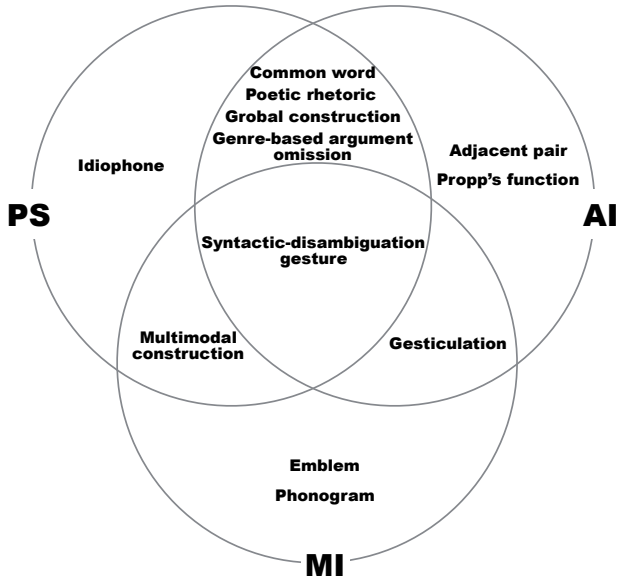

Figure.6 The overall picture of context-containing constructions

wa, as an example of a linguistic element which is both retrospective and prospective.

[4.] Note that phonemic sequence includes mental representations of both speech and writing, because it is produced in our brain and not equal to the actual sound.

[5.] The term *accessible information* is borrowed from Chafe (1994).

[6.] For the sake of convenience, output sets are assumed to comprise a single element in this paper. The single element of output sets can, however, be reduced to some mental representations actually.

[7.] One might object that we can get a particular conceptualization from a single word without any context. Such a situation can, however, be regarded as the context that there is no specific context. My model assumes the network between constructions based on schematicity as Langacker's model. We acquire word constructions which has schematic AI through exposure to specific uses of the word as the same meaning in various contexts (i.e. AI). The more schematic AI a construction has, the more compatible it is with an unfamiliar context. This is why the construction which is invoked without a specific context has usually the most frequent meaning.

[8.] Note that the fact that a construction does not include particular AI does not mean that the AI

contradicts the construction. This is why the construction (8b) can be employed by the mother, who knows that the utterance is said at the 'BEGINNING OF THE TOPIC.'

[9.] In this case, the mother's facial expression or speech characteristics do not play a role of a trigger to invoke the meaning 'REQUEST FOR INFORMATION' unless they are interpreted as 'SHE KNOWS WHO IS GOING TO THAT MEETING' Hence, the construction in (8a) do not include MI.

[10.] Such a figure as Figure 2 can be depicted in every example in this paper. However, they are omitted here for brevity as they would be extremely complicated.

[11.] In the case of *veni, veid, vici,* rhyme construction may not be invoked because the rhyme /i:/ stands for 'FIRST-PERSON SINGULAR PERFECT ACTIVE' in Latin. To the common people, however, Latin conjugation is not familiar and each word in the message is regarded as a single concept rather than a composite unit. Rhyme can, threfore, be recognized in the message as Jakobson pointed out.

[12.] The red signal is usually understood as meaning STOP. However, when you are on a pedestrian crossing it cannot be interpreted so.

[13.] An actual dog's bark and an actual rooster's crow are, as a matter of course, regarded as MI-constructions too. Actual physical sound is significantly different from idiophone because the former is composed of many prosodic aspects such as pitch, volume, length and tone. For example, when listening a low-pitched dog's bark edited by an audio editor, we would not identify what produces the original sound. By contrast, idiophones can be interpreted regardless of their pitch, length and tone because they are composed of only PS.

[14.] As noted above, mental representations of letters are included in PS. They are abstract outlines of letters and distinct from letters themselves. In the example of the bold [**あ**], the outline of the letter invokes the phoneme /a/, and the bold font conveys loud or stressed sound.

[15.] When seeing people say to each other '*good morning*', we can recognize that the conversation is held in the morning. This interpretation cannot be invoked in the case of '*hello*.'

[16.] The same goes for phonemic sequences. We will discuss this issue in the next subsection.

[17.] In Okahisa (in press), if a participant in the experiment displays the N1 gesture and the N2 gesture in this order, the sequential performance is categorized as the N1+N2 gesture type, for example.

References

Allwood, Jens. 2002. Meaning potentials and context: Some consequence for the analysis of variation in meaning. In Cuyckens, Hubert, Rene Dirven, and Taylor John R. (eds.) *Cognitive*

Approaches to Lexical Semantics 29–65. Berlin and New York: Moulton de Guyter.

Chafe, Wallace L. 1994. *Discourse Consciousness and Times*. Chicago: University of Chicago Press.

Clark, Herbert H. 1983. Making sense of nonce sense. In Flores d'Arcais, Giovanni B., and Robert J. Jarvella (eds.) *The Process of Language Understanding* 297–331. London: Wiley Chichester.

Dirven, René. 2005. Major strands in cognitive linguistics. In de Mendoza Ibáñez, Francisco J. Ruiz, and M. Sandra Peña Cervel (eds.) *Cognitive Linguistics: Internal Dynamics and Interdisciplinary Interaction* 17–68. Berlin and New York: Mouton de Gruyter.

Jakobson, Roman. 1960. Closing statement: Linguistics and poetics. In *Style in Language* 350–377. Cambridge: The MIT Press.

Kendon, Adam. 1988. How gestures can become like words. In Poyatos, Fernando (ed.) *Cross-Cultural Perspectives in Nonverbal Communication* Vol.1: 131–141. Toront/Göttingen: C.J. Hogrefe.

Langacker, Ronald W. 1987. *Foundations of Cognitive Grammar, vol. 1: Theoretical Prerequisites*. Stanford: Stanford University Press.

Langacker, Ronald W. 2000. A dynamic usage-based model. In Barlow, Michael, and Suzanne Kemmer (eds.) *Usage-Based Models of Language* 1–63. California: CSLI publications.

Langacker, Ronald W. 2001. Discourse in cognitive grammar. *Cognitive Linguistics* Vol.12: 143–188.

Langacker, Ronald W. 2008. *Cognitive Grammar: A Basic Introduction*. Oxford: Oxford University Press.

McNeill, David. 2000. *Language and Gesture*. Cambridge: Cambridge University Press.

McNeill, David. 2005. *Gesture and Thought*. Chicago: University of Chicago Press.

Okahisa, Taro. in press. Tougoteki aimai hyougen o dentatsu suru sai no washa no shiten: Bunpou kouzou to jesuchaa no kankeisei (What do speakers focus on in uttering syntactic ambiguous expressions: The relationship between grammatical structure and gesture). In *Proceedings of the 19th Conference of the Pragmatics Society of Japan* , The Pragmatics Society of Japan.

Propp, Vladimir Y. 1968. *Morphology of the Folktale*. Second Edition. Austin: University of Texas Press. (Translated by Laurence Scott, revised and edited with a preface by Louis A. Wagner; First Edition published in 1928).

Pustejovsky, James. 1995. *The Generative Lexicon*. Cambridge: MIT press.

Ruppenhofer, Josef, and Laura A. Michaelis. 2010. A constructional account of genre-based argument omissions. *Constructions and Frames* Vol.2: 158–184.

Schegloff, Emanuel A, and Harvey Sacks. 1973. Opening up closings. *Semiotica* Vol.8(4): 289–327.

Steen, Francis, and Mark B. Turner. 2013. Multimodal construction grammar. In Borkent, Michael, Barbara Dancygier, and Jennifer Hinnell (eds.) *Language and the Creative Mind* 255–274. Stanford and CA: CSLI Publications/ University of Chicago Press.

Taylor, John R. 2012. *The Mental Corpus: How Language Is Represented in the Mind*. Oxford: Oxford University Press.

Terasaki, Alene Kiku. 2004 [1976]. Pre-announcement sequences in conversation. In Lerner, Gene H. (ed.) *Conversation Analysis: Studies from the first generation* 171–223. Amsterdam: John Benjamins.

Yamanashi, Masa-aki. 2009. *Ninchi Kobun Ron (Cognitive theory of constructions)*. Tokyo: Taishukan Press.

Zima, Elisabeth. 2014. English multimodal motion constructions. A construction grammar perspective. *Studies van de BKL —Travaux du CBL— Papers of the LSB* Vol.8:. (http://uahost.uantwerpen.be/linguist/SBKL/sbkl2013/Zim2013.pdf)

状況包含構文

人はなぜ 1 つの表現を無限に解釈できるのか

岡久 太郎

本稿では、認知言語学において仮定されている形式と意味の対である構文の適用範囲を拡大し、構文の形式が音素列 (PS)、マルチモーダル情報 (MI)、アクセス可能情報 (AI) の 3 つのカテゴリーから構成されうる「状況包含構文」を想定することで、多義性という考えを排し、これまで様々な分野で個別に説明されてきた現象を統一的な観点から説明することを試みる。これまで、語の多義性に関する議論においては、当該の語の意味は一つであるとして、語用論的ルールを仮定し、語の意味の差異を説明する立場と語の複義性自体を認める立場の 2 つがあった。しかし、いずれの立場においても、語用論的ルールのリスト、ないしは語義のリストが無限に必要となる「無限リスト問題」(infinite-list problem) が生じると言える。そこで、本稿では Taylor (2012) の議論を拡張させ、コンテクストが構文の形式に含まれた状況包含構文を想定し、一見多義的に捉えられる言語単位はより大きな構文の一部でしかないことを示す。さらに、状況包含構文の形式を 3 つに区分することにより、これまで様々な分野で研究されてきた記号的現象が、(1) PS 構文、(2) MI 構文、(3) AI 構文、(4) PS-MI 構文、(5) PS-AI 構文、(6) MI-AI 構文、(7) PS-MI-AI 構文の 7 つの観点から統一的に捉えられることを示す。この考え方を採用することで、これまでは、意味論、語用論、詩学、文字論、物語論、マルチモーダル分析等において研究されてきた様々な記号現象を基本的で単純な情報に還元して分析することが可能となり、記号の理解／産出メカニズムの解明に繋がることが期待される。