MSFA を用いた文章の深い意味記述

†独立行政法人 情報通信研究機構 けいはんな情報通信融合研究センター E-mail: †{yshibuya,jhlee,kuroda,isahara}@nict.go.jp

あらまし キーワード

Specifying deeper semantics of a text using MSFA

Yoshikata SHIBUYA[†], Jae-Ho LEE[†], Kow KURODA[†], and Hitoshi ISAHARA[†]

† National Institute of Information and Communications Technology (NICT) Knowledge Creation Research Center 3-5 Hikaridai, Seika-cho, Souraku-gun, Kyoto, 619–0289 Japan E-mail: †{yshibuya,jhlee,kuroda,isahara}@nict.go.jp

Abstract In this article, we introduce the JCASR Project (in progress at NICT), which aims to develop a relatively small Japanese corpus of texts annotated for "semantic frames" and their "frame elements" (aka "semantic roles"), using *Multilay-ered/Multidimensional Semantic Frame Analysis* (MSFA) [3]. MSFA is a framework of semantic annotation/analysis compatible with the Berkeley FrameNet project [1], [6] that provides a multidimensional description of "contextualized" meanings of words and phrases. After outlining the project, we provide case studies of two sentences (one from a newspaper article and the other of a discourse).

Key words contextualized meanings of words and phrases, FrameNet, the JCASR project, Multilayered/dimensional Semantic Frame Analysis (MSFA), semantic role tagging

1. Introduction

The past 20 years have seen the maturation of surface-true, distributionally and statistically based knowledge acquisition techniques. The development has offered us a parse with a good coverage, free from inconsistencies immanent in manual analysis. It was like a new age of NLP. It turns out, however, that such methods clearly have limitations. One of the critical problems is that these techniques do not always meet our need for processing "deeper" semantics (or "shallow pragmatics") in and of the language people use everyday. What we mean by "deeper semantics" is a class of semantic specifications that goes beyond simple "word senses/meanings" including the so-called "inferences" (yet applying this term would be no help here, because nobody knows exactly what "inferences" are, after all). This is the classical problem of Knowledge Representation (KR). Suppose that the KR problem is resolved totally after we have finished building all the relevant ontologies and inference engines running over them. Did we get to our ultimate goal at last? We suspect the answer is no, because we still need to find out how pieces of knowledge are linked to pieces of language. This is what Kuroda and Isahara (2005) called the problem of Knowledge-Language (K-L) Linking.

If our suspicion is true, we then need to construct a wide-coverage database that specifies what kind of linguistic units are linked to what pieces of world knowledge, or "ontologies" without (too simplistically) assuming that words denote "concepts" (i.e., building blocks of ontologies). This assumption is far from true, however: virtually every word of a given language is ambiguous, and it is far from well-understood how ambiguity comes into play, although there are good theories of word sense disambiguation/creation like Generative Lexicon (Pustejovsky [?]). This reality makes it very ineffective to state that words "denote" concepts. In fact, this situation demands us to build lexical resources dedicated to the identification and specification of "units" of the K-L Linking.

1.1 The JCASR project

Development of a *Japanese Corpus Annotated for Semantic Roles* (JCASR) is being attempted as one of the research projects at the National Institute of Information and Communications Technology (NICT), Japan. The project is proceeded with a crucial assumption that "units" of the K-L linking are "idealized/(stereo)typical situations" which can be identified as (semantic) frames in the same sense as the Berkeley FrameNet (BFN) project (Baker et al. [1]; Fillmore et al. [2]). The goal of the JCASR project is to construct a (relatively small) corpus of Japanese texts annotated for (seman-

tic) frames and their frame elements (aka "(situationally-defined) semantic roles"). The goal is to establish a set of (ontological) links from "pieces of world knowledge" to text segments in terms of semantic role tagging.

1.2 What we mean by "semantic role tagging"

In our approach, a strong emphasis is given to the identification and specification of finer-grained, situation-specific roles at concrete levels (e.g., $\langle Robbers \rangle$, $\langle Victim \rangle$, $\langle Valuables \rangle$; hereafter, $\langle R \rangle$ represents a role) rather than those of coarse-grained, general-purpose roles at abstract levels (e.g., $\langle Agent \rangle$, $\langle Patient \rangle$, $\langle Theme \rangle$). Section 3 is devoted to illustration of this approach. Our approach is theoretically motivated by the hypothesis we assume that deeper, and "better," understandings are achieved at more concrete levels, rather than at more abstract levels. This hypothesis is one of the points that differentiate MSFA from other (usually more "formally oriented") approaches to semantic annotation/analysis which tend to assume that the deepest semantic analysis is the most abstract semantic analysis. More formally, we posit the following:

Concreteness bias on semantic interpretation:

the more specific and concrete your understanding is, the better it is (as long as it is not obviously wrong).

A lot of phenomena suggest that "deep enough" semantic analysis of a text demands effective specifications of what guesses people make, as well as of semantic types of text segments. What is suggested is that it does not really matter whether people's understandings are semantically based or pragmatically based as far as our goal is to illustrate people's text understanding: to specify what people understand is the point at issue, while how they do so is not. The semantics/pragmatics distinction would make sense if the issue is how people understand (after we have clarified what they understand). This reasoning would be both good news and bad news, depending on your perspective. It is good news if you feel that routes to deeper semantics are promised. It can also be considered bad news if you feel that you cannot be so optimistic as to say "Leave it all to pragmatics" any more, because what is at issue now is what pragmatics does and how it works: you need to specify it.

2. Development Scheme

2.1 Status of the project

The JCASR project officially began two years ago. It is (still) at a preliminary, "exploratory" stage. At the moment, we are trying to find out what kinds of frames/situations are found at what granularity levels without assuming a pre-existing, "ready-to-use" database of semantic frames and their frame elements. Some preliminary results have been reported in Kuroda et al. (2006), for instance. We have not started serious development of a semantically tagged corpus yet, but annotation samples are available freely or privately at the web sites (contact us for more details). It should probably be noted that we are currently working independently of the Japanese FrameNet (JFN) project (Ohara et al. 2003). But we are also negotiating with the BFN staff to make the MSFA-based annotation scheme shared with BFN.

2.2 Overview

Currently, we are following the "incremental" development scheme as per the following: (1) select a Japanese text T from a text

database; (2) have each sentence of T segmented into text segments by the staff at NICT (each result of segmentation always needs to be checked manually, because the standard outputs of the so-called "morphological analyzers" like KNP and ChaSen are sometimes inappropriate for our purposes; this issue to be mentioned again later below); (3) ask "external" annotators to annotate the segmented texts by making reference to databases D_1 and D_2 of "sample annotations" hosted at the web sites (available both publicly and privately); (4) collect the annotations conducted by annotators as "drafts," and check and edit the results if necessary, which is very often the case (this is conducted by the staff of the project group at NICT); (5) add the edited results to the databases D_1 and D_2 , and "sanitize" the databases when needed.

T is always chosen from Japanese texts which are aligned with English texts. This is because we expect that future comparisons against other annotations (using the BFN database, for example) can be facilitated. So far, all texts have been taken from the following text bases:

 D_1 : Sample annotations for texts from a collection of English-Japanese alignments of copyright-free texts like Aesop's Fables are hosted at [http://www.kotonoba.net/mutiyama/cgibin/hiki/hiki.cgi?FrontPage]. No access restriction applies.

D₂: Samples for texts from *Kyoto University Corpus* are hosted at [http://www.kotonoba.net/~mutiyama/cgi-bin/hiki2/hiki.cgi?FrontPage]. Access restriction applies.

The original texts for D1 are provided at [http://www2.nict.go.jp/x/x161/members/~mutiyama/align/index.html]. Tentatively, we separate the procedures to identify (a) frames for event conceptualizations (e.g., $\langle Robbery \rangle$, $\langle Predation \rangle$) and (b) frames for social interactions (e.g., speech acts like $\langle Claiming \rangle$, $\langle Criticizing \rangle$, $\langle Doubting \rangle$, $\langle Protesting \rangle$, $\langle Warning \rangle$). This is because the second type of frames are more complex, more data selective, and harder to specify. Currently, only Hajime Nozawa is working on the second type (see § 3.3). Nozawa's work has not yet been integrated into the results of the first type worked on by Kuroda, Lee and Shibuya.

3. Case Studies

3.1 The procedure

In conducting an MSFA, one employs a table T of m+3 rows and n columns: m is the number of text segments (including "null instantiations"), and *n* is the number of frames identified as comprising the "understood content" of a sentence s. The cell at (i, j) of T specifies the semantic role r of the jth frame fj. The value for fj.r includes "null," which means "non-realized role for f." As illustrated in the following sections, the MSFA's analytical scheme includes three header lines: "Frame ID" (row 1), "Frame-to-Frame relations" (row 2), and "Frame Name" (row 3). After the completion of the segmentation task, one turns to fill in the cells rightward, specifying (or identifying) Frame Names (together with the names of Frame Elements) and Frame-to-Frame relations among these frames. Frame IDs are local variables used to specify Frame-to-Frame Relations, whereas Frame Names and Frame-to-Frame Relations are global variables. Each sentence of a text T is segmented into text segments before one starts annotating them. The sentence segmentation process is conducted by using morphological analyzers such as ChaSen

Frame ID	F1	F2	F3	F11	F8	G1	F4	E5	F6	F7	F9	F10	F12	F13
	FI		F3	FII		G1		F5	presupposes	F7	F9	elaborates		F13
Frame-to-Frame	elaborates F2	elaborates F3; realilzes F1	realizes F2	elaborates F8	constitutes F3; realizes G1		presupposes F5;		F4,F5; presumes	constitutes F5	elaborates F5	F9,F12;	presupposes	
Relations		realilzes F I			realizes G1		constitutes F8		F4			presupposes F5	F13; realizes F13	
				~過去性の指定	40.04 71.4	~値の指定[役割				(W.cla		将棋[プロ の]//Shogi		
Frame name	報道//News	報告//Report	語り//Narration	~//~Tense	~特徴づけ ~//~Characteriza	の]~//~Value	挑戰//Challenge	勝負//Match	勝利//Victory	優劣 //Advantages/Dis	将棋//Shogi	(Japanese	職(業)をもつ	生計立て//Making
Frame name	報題//News	₩ = // Keport	語9//Narracion	specification[for	tion~	specification[ofor	964X// Challenge	mp更//Match	勝行// VICTORY	advantages/ bis	(Japanese chess)	chess)[profession	//Occupation	a living
				event]~	tion-	role]~				auvantages		all		
*	報道者 //News	報告者 //Reporter	語り手 //Narrator	指定者 //Specifier	特徴づける者									
	reporter 受報者 //News	受報者 //Report		70.7C U 11 Option 11	//Characterizant									
*	receiver	receiver	聞き手 //Listener											
	報道内容//	報告内容										プロ将棋.EVO: タイ	内容[職業	生計を立てる手段
*	Content of news	//Content of	内容 //Content									トル//Professional shogi EVO: title	の]//Content of	//Means of making
		report						場所[勝負の]//				snodi.EVU: title	occupation	a living
*								Place for match						
				~過去性の指定[事 能		役		時点[勝傷						
昨年//last year				類 の]~.GOV[1,2]//	対象.Attr	割.Attr//Role.Att	時間//Time	の]//Time of	時期//Time					
*F477last year				~Tense	//Object.Attr	r r	wyjaj// Tittle	match	Myses// Tille					
				specification[for										
							時	時点[勝負	時					
.//,				EXT			間.EXT//Time.EXT	の].EXT//Time of match.EXT	期.EXT//Time.EXT					
								match.EX1				プロ将棋.EVO: 棋士		手段により生計を立
¥ = / O/				参与者[事態 の][1]//Event			#PT //O	勝負者[1]	86 (4b) #r (/ 1	優位者//Person	棋士[1,2]//Shogi		職(業)をもつ者	てる者[1]//Person
米長//Yonenaga				Participant[1]			相手 //Opponent	//Player[1]	敗(北)者// Loser	with advantage	player[1,2]	al shogi.EVO:	[1]//Person in employment[1]	using means to
IZ//to				r ai dcipant[i j			MARKER//NA					player[1.2]	employment[1]	make a living[1]
IC//t0										優劣.EVO				
挑ん//challenge				事態//Event			挑戦.GOV//	勝負.EVO	勝利.EVO//	//Advantages/Dis				
							Challenge.GOV	//Match.EVO	Victory.EV0	advantages,EVO				
				~過去性の指定[事 能										
t€//~PAST~				の]~.GOV[2,2]//			EXT							
70771710111				~Tense			Di.							
				specification[for										
の//~N.A.~					MARKER ~特徴づけ	役割//Role			-			-		
					~.GOV[1,2]//~C									
が//~NOM~					haracterization~.	MARKER								
					GOV[1.2]									
				参与者[事態	44.04		All Market / /	WW #107		ds (4 * / / D	## 1 52 23 / /C' ·	プロ将棋.EVO: 棋士 [2,2]//Profession	職(業)をもつ者	手段により生計を立
羽生//Habu				の][2]//Event	特徴 //Characteristics	值//Value	挑戦者// Challenger	勝負者[2] //Player[2]	勝(利)者//Winner	劣位者// Person with disadvantage	棋士[2,2]//Shogi player[2,2]	al shogi.EVO:	[2]//Person in	てる者[2]//Person using means to
				Participant[2]	// CridiaCteristics		Challenger	//riayer[2]		with disadvantage	player[2,2]	player[2,2]	employment[2]	make a living[2]
					~特徴づけ									
**					~.GOV[2,2]//~C									
					haracterization~.									
.//.	FXT	FXT	FYT		GOV[2,2] FXT									
077.	LAI	LA	LΛI		LAI		1		1			1		

図 1 An MSFA of (1)

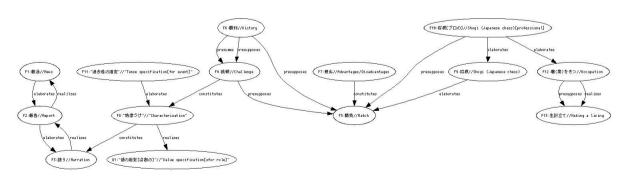


図 2 An SFNA of (1)

(http://chasen.naist.jp/hiki/ChaSen) and KNP (http:// nlp.kuee.kyoto-u.ac.jp/nl-resource/knp.html). It is important to note that the segmentation is also conducted manually by the annotators of the project group. This is because for our purposes it is often the case that the outputs of the morphological analyzers leave something to be modified at times (recall §2.1). Another reason is that in an MSFA one tries to specify as many links as possible between knowledge pieces to often discontinuous, multiword units like idioms and proverbs. In general, MSFA does not assume the principle of compositionality, in that for the phrase p of $w_1 + \cdots + w_i + \cdots + w_n$, the meaning of p is strategically regarded as independent of w_i . This is true not only of opaque units like idiom chunks but also of virtually any units which are usually considered to have "regular" semantics. This decision is necessary to avoid the petitio principii in semantic specification, even if it seems redundant.

3.2 Sample 1: A segment of a newspaper article

Consider the sentence in (1) taken from *Kyoto Universtiy Corpus* (Kurohashi and Nagao [5]). (2) gives the English translation. Table 1 provides an MSFA of (1):

- (1) XXXXXXXXXXXXXXXXXX [S-ID:950101075-033]
- (2) Last year, Habu challenged Yonenaga for the title.

In Table 1, the annotations have been converted from Japanese to English (the frame names are given tentatively in English; for the original MSFA in Japanese, see http://www.kotonoba.net/ ~mutiyama/cgi-bin/hiki2/hiki.cgi?c=view&p=msfa-round2-All but the first column represent the frames identified, or "discovered" in the analysis. As shown in the table, 14 frames were identified in (1) (but note that the number is subject to further modifications; see the later discussion). Each of the frames consists of some semantic roles. The $\langle Match \rangle$ frame (= $\langle XX \rangle$ in the original Japanese analysis; the URL given above), for example, includes the roles such as $\langle Player \rangle$ (= $\langle XXXX \rangle$), $\langle Place for match \rangle$ (= $\langle XX[XXXX] \rangle$), $\langle Time \rangle$ of match \rangle (= $\langle XX[XXXX]\rangle$). The colored cells indicate the semantic roles that are considered to be realizing their values in the relevant frames. The uncolored (or "empty") ones, on the other hand, represent those that do not seem to have any specific roles in the frames. In conducting an MSFA, it is the semantic roles not the semantic types that are annotated in analyzing a text. Semantic types are roughly equated with natural kinds. In contrast, semantic roles are "situation-specific concepts" which are considered highly cultureparticular and hence are taken to play a more crucially important role than semantic types in one's understanding of a text (Kuroda et al. [4] [revised]; for a related discussion, see also Kuroda and Isahara [3]). The table contains *-symbol, which represents elements that are not lexically realized in a text. For example, the $\langle \text{News} \rangle$ frame (= $\langle XX \rangle$) and the $\langle \text{Report} \rangle$ frame (= $\langle XX \rangle$) contain some semantic roles in the rows with *-symbol. This is because given these frames such semantic roles must be specified as well. Having outlined the MSFA table for (1), one might address an important question: How deep should one go in describing the meaning of a sentence? In the current MSFA scheme, a sentence is typically given over 20 frames, and each of which comes with some semantic roles as illustrated above. To compare the MSFA scheme to another semantic-annotation framework, the BFN, for example, customar-

ily limits the number of frames to 2 or 3 frames per sentence. It is remarkable that in an MSFA a sentence as short as (1) is identified with a rather large number of frames compared to the BFN scheme (14 frames in the case of 1: recall Table 1). The deep semantic description as illustrated in Table 1 is the result of meeting one of the fundamental theoretical principles of MSFA: the so-called "Be greedy" principle (Kuroda et al. ?? [revised]), which reads as follows: the analysis/annotation needs to be greedy, in that as many frames as you need can be identified and added to the analysis, as long as they are found necessary for providing deep enough semantic analysis of a text. From the MSFA perspective, sample semantic annotations provided by the BFN framework today are not taken to be deep enough. It is important to note that we do not claim the MSFA illustrated in the table to be as deep a semantic description as we can provide. Rather, it should be taken as exemplifying a "tentatively" suggested version of our MSFA of (1). The position that MSFA maintains as to the so-called "granularity" problem is that it must be worked on inductively. The view we hold is that adequate levels of granularity need to be "discovered" through inductive exploration into real texts. The Frame-to-Frame relations are described in row 2 of Table 1, which is depicted in Figure 1 (called a Semantic Frame Network Analysis [SFNA]). In the figure, each frame identified in the MSFA (Table 1) is represented in a circle and the arrows indicate how these frames are interrelated with one another. The diagram is generated automatically by Graphviz based south specification given in Table 1 (i.e., the MSFA of (1)). Note that here, as well as in Table 1, all the frame names have been translated from Japanese to English. It should be noted that the Frameto-Frame relations are not assumed a priori in MSFA. Instead, we hold that the set of such relations too needs to be discovered by inductive processes, rather than in a top-down, theory-driven manner. Enlisted in (3) below are the relationships (the specifications partially omitted for lack of space) that are often found. The list is not exhaustive. Most of the Frame-to-Frame relations in the BFN are analogous to those of the MSFA; but for the discrepancies, see Kuroda et al. [4] [revised]: §??).

(3) (a) "Elaboration" relation: A frame F elaborates another frame G; i.e., F inherits from G. (b) "Constitution" relation: F constitutes G; i.e., F is part of G. (c) "Presupposition" relation: G presupposes F. (d) "Presumption" relation: F presumes G. (e) "Realization" relation: F realizes G. (f) "Target/Transfer" relation: F targets G.

Having provided a brief illustration of how one analyzes a newspaper article sentence of (1) with MSFA, let us now turn to the prospects of this framework for cross-linguistic semantic annotation research. Consider (4), which is the Korean translation of (1) (translation provided by Jae-Ho Lee):

- (4) Jagnyeon -e yonenaga -ege dojeonha n salam -i habu.
- An MSFA of (4) (here omitted for lack of space) reveals the similarities and differences in Japanese and Korean in terms of the availability of the frames (in 1 and 4). Below is the list of the additional frames that were identified with the Korean sentence (4):
- $(5) \quad (i) \ensuremath{\left\langle \sim \text{Specification[of being Japanese]} \right\rangle} (=\ensuremath{\left\langle \sim \text{XX[XXXX]} \right\rangle}) \\ (ii) \ensuremath{\left\langle \sim \text{Specification[of Time]} \right\rangle} (=\ensuremath{\left\langle \sim \text{XX[XXXXX]} \right\rangle}) (iii) \ensuremath{\left\langle \sim \text{Specification[of difference in ability]} \right\rangle} (=\ensuremath{\left\langle \sim \text{XX[XXXXXXX]} \right\rangle}) (v) \ensuremath{\left\langle \sim \text{Modification} \right\rangle} (=\ensuremath{\left\langle \sim \text{XX}[\text{XXXX]} \right\rangle}) (v) \ensuremath{\left\langle \sim \text{Modification} \right\rangle} (=\ensuremath{\left\langle \sim \text{XX} \right\rangle}) \\ \end{aligned}$

(vi) $\langle \sim$ Specification[of the fact that it is an interpersonal event] $\sim \rangle$ (= $\langle \sim$ XX[XXXX] $\sim \rangle$)

The involvement of these Korean-specific frames is considered to be due to the syntactic characteristics of the Korean language. It is interesting to see that Korean and Japanese (two typologically close languages) differ in the availability of the frame types in understanding an equivalent sentence. The results of an MSFA of (4) suggest an interesting prospect for cross-linguistic (semantic annotation) research, because it is suggested that a careful semantic analysis using MSFA would make a contribution to clarifying the universals and particulars of languages with respect to how people understand a sentence.

3.3 Sample 2: A segment of a prose

Let us briefly discuss how one analyzes a sentence from a discourse with MSFA. Consider the sentence in (6), taken from Aesop's *Fables* (*The Ass and the Grasshopper*):

(7) They replied, "The dew." [http://www2.nict.go.jp/x/x161members/mutiyama/align/htmPages/aesop-0.htm]

An MSFA of (6) is given in Table 3 and an SFNA in Figure 2. Similar to Table 1 and Figure 2, all the annotations of Table 3 and Figure 4 have been converted from Japanese to English (for the original analyses, see http://www.kotonoba.net/~mutiyama/ cgi-bin/hiki/hiki.cgi?c=view&p=msfa-aesop01-s03-s04). A total of 25 frames were identified with (6). A remarkable fact with (6) is the involvement of the frames such as $\langle Narration \rangle (=\langle XX \rangle)$, $\langle \text{Telling a story} \rangle = \langle XX \rangle$, $\langle \text{Creative work} \rangle = \langle XX \rangle$, $\langle \text{Allegory} \rangle$ creation \rangle (= $\langle XXXXX \rangle$), $\langle Imagination \rangle$ (= $\langle XX \rangle$), which are identified because of the nature of the text (i.e., a discourse-style fable). Of the frames found in the MSFA of (6), the (Having curiosity) frame (=\langle XXXXXXXX\rangle) is perhaps among those that significantly differentiate MSFA from other semantic annotation frameworks for its elaborate semantic specification. In MSFA, one recognizes the (Having curiosity) frame in a reply sentence like (6), because it is a sentence with which the grasshoppers (specified by they) answer the preceding question that was posed by the ass who was "curious" to find out what type of food the grasshoppers live on so that they could possess such beautiful voices (see the preceding context given below; also notice the description given in Table 2 that says "G6 motivates F9"):

(8) XXXXXX XXXXXX X X XXX XXX XXXXXXX

(9) AN ASS having heard some Grasshoppers chirping, was highly enchanted; and, desiring to possess the same charms of melody, demanded what sort of food they lived on to give them such beautiful voices.

In this section, we have briefly introduced how one analyzes a discourse segment, showing that MSFA provides a framework for describing discourse understanding as well. There is, however, a caveat to note before closing this section. The caveat to bear in mind is that the current version of MSFA scheme has not yet been equipped with a descriptive tool to link a reply to the descriptive contents of the preceding sentences effectively (one possible scheme is being developed by H. Nozawa; see his paper in this volume).

4. Conclusion

In this article, we introduced the JCASR Project, which aims to develop a relatively small Japanese corpus of texts annotated for semantic frames and their frame elements (aka semantic roles) in the same sense as the Berkeley FrameNet. After outlining the project, including its status, methodological procedures, strategies, and so forth, we provided case studies of two sentences, suggesting that MSFA provides a framework for deep semantic description.

参照文献

- C. F. Baker, C. J. Fillmore, and J. B. Lowe. The Berkeley FrameNet project. In *COLING-ACL 98, Montreal, Canada*, pp. 86–90. Association for the Computational Linguistics, 1998.
- [2] C. J. Fillmore, C. R. Johnson, and M. R. L. Petruck, "Background to FrameNet," *International J. of Lexicography*, 16 (3), pp. 235–250, 2003.
- [3] K. Kuroda and H. Isahara, "Proposing the Multilayered Semantic Frame Analysis of Text," Proc. of the 3rd International Conference on Generative Approaches to the Lexicon, pp. 124-133. Geneva, Switzerland, 2005. [http://clsl.hi.h.kyoto-u.ac.jp/~kkuroda/papers/msfa-gal05-rev1.pdf].
- [4] K. Kuroda, M. Utiyama, and H. Isahara, "Getting deeper semantics than Berkeley FrameNet with MSFA," Proc. of the LREC-06 (P26-EW), Genova, Italy, 2006. [Available at: http://clsl.hi.h.kyoto-u.ac.jp/~kkuroda/papers/msfa-lrec06-rev1.pdf]
- [5] S. Kurohashi and M. Nagao, "Building a Japanese parsed corpus: While improving the parsing system," in *TreeBanks*, ed. A. Abeill'e, pp. 249–260, Kluwer, Dordrecht, 2003.
- [6] J. B. Lowe, C. F. Baker, and C. J. Fillmore, "A frame-semantic approach to semantic annotation," *Proc. of the SIGLEX Workshop on Tagging Text with Lexical Semantics*, pp. 18–24, Washington, D.C., 1997.
- [7] K. H. Ohara, S. Fujii, H. Sato, S. Ishizaki, T. Ohori, and R. Suzuki, "The Japanese FrameNet Project: A Preliminary Report," *Proc. of PACLING* '03, pp. 249–254, 2003.
- [8] J. Pustejovsky, The Generative Lexicon. MIT Press, Cambridge, MA. 1995.

Frame ID	F8	F10	G3	G4	G7	F7b	F7a	Н3	F9	G5	G6	F13	F12	F11	G2	F6	H1	H2
Frame ID	elaborates F5;	FIU	63	64	- 47	F/D	F/8	elaborates F7:	presumes F8;	65	GD	FIS	constitutes		62	- 10		n ₂
Frame-to-	presuposes	constitutes				constitutes	constitutes	constitutes	elaborates		motivates		H2;			constitutes	presupposes	
Frame	F9; elaborats		elaborates G4			F10; realizes	F10; realizes F9;	G4;	F7a;	constitutes	F9;		constitutes		constitutes	J1;	F4; presumes	constitutes H1
Relations	H3;	G3				F8,H3;	elaborates G7	presupposes	constitutes	G4	?motivates		H3;		H2	presupposes	H2	
	constitutes					elaborates G7		G5	F10; realizes		G5		presupposes			F7b		
	F10								G5				F13 ~値の指定[役		~複数性の指			
			会話		行為[意図	発言		提供[情報			好奇心をもつ	~特徴づけ	割		定[対象		擬人化	~特徴づけ[人間
Frame Name	返答//Reply	対話	//Conversatio	共同行為	001//Actionfu		発言	が //Offer [of	質問	要求	対可心をもう //Having	~//~Charact	の]~//~Valu	摂食//Intake	の]~//~	引用	//Personificati	の]~//~Charac
Frame Name	AEM//Reply	//Dialogue	n	//Joint Action	olitional]	2]	[1]//Remark[1]	information1	//Question	//Request	curiosity	erization~	e	10136//IIICake	Plurality	//Quotation	on	terization[hum
					Ontional	-1		miorination			curiosity	CHEUGOT-	specification		specification		on.	an]~
												-	[for role]		[for object]			
																		行特徴づけ者
*															指定者	引用者	擬人化者	//Characteriza
															//Specifier	//Quoter	//Personifier	nt
*																		
			l						l	1				l		l		
			1		1			1	1			i e					擬人化の先	
			l						l	1				l		l	[+human][2]	特徴[人間の][2]
*																	Target for	Characteristics
																	personification	[of human][2]
																	[_hhuman][2] 擬人化の先	
			l						l	1				l		l	[+human][1]/	特徴[人間
*																	/Target for	の][1]//Charac teristics[of
																	personification	human][1]
	ata sov Care pp	At STAR Y DE DE	先行する話題				内容[質問の]//		内容[質問の]//								[#human][1]	numanjji
*	内容[質問 の]//Content[先行する話題 //Preceding	先行する話題 //Preceding				内容[実同の]// Content[of		内容[資同の]// Content[of									
	of question1	tonic	tonic				question		question1									
			会話者	共同行為者[1]	行為者[意図												擬人化の元	
	返答をもらう者	対話者	[1]//Convers	//Joint action		相手[発言の]//	発言者	受取り手	質問者	要求者//	好奇心をもつ					相手[発言	[2]//Source of	
*	// Person who	[1]//Interlocu	ation	participant	Agent[volition	Listener[of	//Remarker	//Receiver	//Questioner	Requester	者//Person					の]//Listener[personification	
	gets reply	tor[1]	participant[1]	[1]	al][1]	remark]					with curiosity	1				of remark]	[2]	
		4100.00	会話者		行為者[意図		40 F (TO TO TO TO TO		10 IC C 0000 - 1 / /	相手[要求	帰属体[好奇心	特徴づけの対					擬人化の元	
キリギリス	返答者	対話者 [2]//Interlocu	[2]//Convers	共同行為者[2] //Joint action	Mh1[21//	発言者	相手[発言の] //Listener[of	提示者	相手[質問の]// Questionee[of	の]//	の対	象//Object	指定者	摂食者	対象//Object	発言者	[1]//Source of	
//grasshopper	//Replier	tor[2]	ation	participant[2]	Agent[volition	//Remarker	remark1	//Offerer	questionee[or	Requestee[of	象]//Source	of	//Specifier	//Intaker	x) #//Object	//Remarker	personification	
		tortel	narticinant[2] 会話者	共同行為者	al1[21		remarkj		questionj	request1	of curiosity 帰属体[好奇心	characterizat 特徴づけの対					F11	
		対話者	[2].EXT//Con	[2].EXT//Join	行為者[意図	発言		提示	相手[質問	相手[要求	の対	象.EXT//Obje		摂食		発言		
たち	返答者EXT//	[2].EXT//Inte	versation	t action	B)][2].EX1//	者.EXT//Rema		雅.EXT//Offer	の].EXT//	の].EXT//	象1.EXT//Sou	ct of	指定者.EXT//	新.EXT//Intak	GOV	希.EXT//Rema		
//~PLURAL~	Replier.EXT	rlocutor[2].EX	participant[2]	participant[2]	Agent[volition	rker.EXT		er.EXT	Questionee[of		rce of	characterizat	Specifier.EXT	er.EXT		rker.EXT		
		T	FYT	FYT	al][2].EXT				question].EXT	request].EXT	curiosity	ion FXT						
t‡//~TOPIC~	MARKER	MARKER 対	MARKER 会話.EVO//		1		-	ļ				-	-	l	-	MARKER		-
答え//reply	質問.GOV//	对 話.EVO//Dialo	受話.EVU// Conversation.						質問.EVO//	要求.EV0//				l		EVO?		
and//reply	Question.GOV	que.EVO	EVO		1				Question.EVO	Request.EVO						LVO		1
た//~PAST~	EXT	EXT	EXT						EXT	EXT						EXT		
.//.									ļ									1
5.400							MARKER[1,2]		l	1				l		ELIOVEDIA OL		
Γ//"	MARKER[1,2]	MARKER[1,2]	MARKER[1,2]			MARKER[1,2]	MARKER[1,2]		l	1				l		EVOKER[1,2]		
	at start terms	ata start name = 2	内容[会話			about Title C	distribution of the		010x100000 m 2 1 1	対象[要求	546-F47 (F. 1	特徵//						
**//it's	内容[返答 の]//Content[内容[対話の] //Content[of	の]//			内容[発言 の]//Content[内容[発言の] //Content[of	情報	対象[質問の]// Object[of	の]//	対象[好奇心 の]//Object	特徴// Characteristi	役割//Role			発言//Remark		
//ILS	of reply]	//Content[or dialogue]	Content[of			of remark]	remark]	//Information	question]	Object[of	of curiosity	cs	1X89//R0le			元名//Remark		
	or reply]	alaloguej	conversation1		 	Ji remark)	remarkj		question	request1	or currosity	C3		対象[摂食の][=				-
												1						1
水滴//dew		_											值//Value	食物] //Object[of				
水滴//dew		_											值//Value	食物」 //Object[of intake][=food				
														//Object[of				
水滴//dew だ//be		_						EVOVER					值//Value	//Object[of				
								EVOKER [acomosite]						//Object[of				
だ//be	MARKER [2,2]	MARKERI 2.21	MARKER[2,2]			MARKER[2,2]	MARKER(2,21	[+composite]						//Object[of		EVOKER[2,21		

図 3 An MSFA of (6)

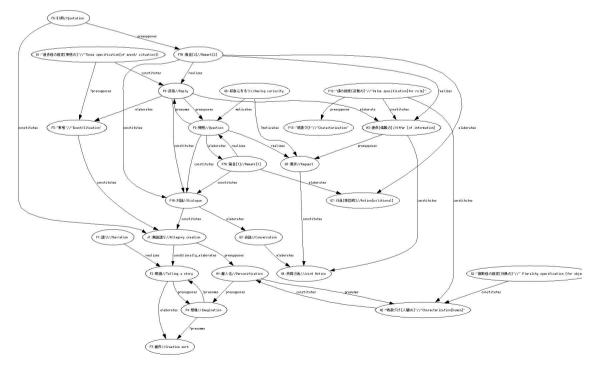


図 4 An SFNA of (6)