Japanese FrameNet: Toward Constructicon Building for Japanese

JAPANESE FRAMENET

Kyoko Hirose Ohara
Keio University, Japan
ohara@hc.st.keio.ac.jp

11th Aug, 2012
ICCG7
Hankuk University of Foreign Studies, Seoul, Korea

Japanese FrameNet project

Toshio Ohori   (The University of Tokyo)
Seiko Fujii   (The University of Tokyo)
Ryoko Suzuki   (Keio University)
Hiroaki Saito   (Keio University)
Hiroaki Sato   (Senshu University)
Shun Ishizaki   (Keio University)

2

Current annotators & programmers

Toshiko Kigoshi
Hiroya Hagino
Naoko Kurokawa
Hidetoshi Kobori
Taisei Suzuki
Dorian Martin
Elise Nguen Hoang

3

Outline

1. Introduction
2. Difficulties in Building Japanese FrameNet
3. Differences in Frames and Constructions between FrameNet and Japanese FrameNet
4. Implications of Lexicon & Constructicon Building for Frame Semantics and CxG
5. Conclusions

4

FrameNet (FN)

- On-line lexical resource for English, based on frame semantics and supported by corpus evidence
  - Frames
    - "[A] script-like conceptual structure that describes a particular type of situation, object, or event along with its participants and props" (Ruppenhofer et al. 2010)
    - Related through **frame-to-frame relations**
    - **Frame Elements**
      - Participant (or prop) roles of the frames are identified and defined
    - Words are grouped based on the frame they evoke
    - A **Lexical Unit (LU)** is the pairing of a word and frame

5

FrameNet (Continued)

  - Lexicon
    - Sentences that represent clearly the meaning and combinatorial possibilities of Lexical Units are collected from a corpus and their respective frame elements are annotated
  - "Constructicon"
    - A registry of English constructions

"[a] **lexicon** should specify the grammatical affordances of its entries; a **grammar** should specify the kinds of lexical units capable of occurring in specifiable positions within grammatical constructions. The most consistent way to represent such mutual dependencies would be to provide both kinds of information in a single well-articulated **grammar + lexicon**" (Fillmore 2006: 35; emphasis in the original)

6

Japanese FrameNet (JFN)

- Creating a **prototype of an on-line Japanese lexical resource** following FrameNet methodology and practice
  - Describes the **sense of each lexical unit** with respect to the **semantic frame** it evokes
  - **Annotates corpus examples** of each word analyzed with **frame elements**
- Compatibility with FrameNet
  - JFN databases and annotation tool
  - JFN frames: imported from FN (the Expand approach)
  - Annotation methods
- Lexicon building > Constructicon Building

7

JFN Lexicographic Annotation

Relation frame
- A relation holds between Entity1 and Entity2.
- Frame Elements: Entity1, Entity2
- [ENTITY1 **DNI**] [ENTITY2 **hito to**]   dono.yooni **tukia**ttara ii …
      people COM how   interact adequate
  '… how shall (I) <u>interact</u> with people?"

8

JFN Full Text Annotation

9

Japanese FrameNet (JFN)

Research questions:

- To what extent is the Frame-semantic approach suitable for analyzing the Japanese lexicon?

- To what extent are the existing English-driven semantic frames applicable to characterizing Japanese lexical units?

10

Difficulties in Building Japanese FrameNet

11

Writing System

LUs which begin with Hiragana

LUs which begin with Katakana
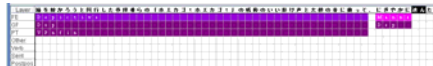
LUs which begin with Chinese Characters

12

Word Boundary Issues

JFN database storage of Japanese sentences
Original Japanese sentence: 彼の家へ行く
In JFN database:　　　　　彼　の　家　へ　行く

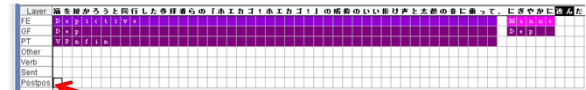FN annotation tool (FNDesktop)
(Lee-Goldman and Rhodes 2009)

JFN annotation tool (JFNDesktop)

13

Postpositions

Postposition Layer

カラ◇	ガ◇	デ◇	ト◇
マデ◇	ノ◇	ハ◇	ヘ◇
	モ◇	ヨリ◇	ヲ◇

Choices for Postposition

14

Postpositions: Default Rules

Default rules concerning combinatorial patterns of FE, GF, PT, and Postpos

　　Postpos=o → PT=NP, GF=Obj
　　　　　　Acc
　　Postpos=ga / no → PT=NP, GF=Ext
　　　　　　Nom Gen
　　Postpos=NULL → PT=AVP, GF=Dep
　　Postpos=kara / de / to / ni / e / made / mo / yori → PT=NP, GF=Dep
　　　　　Abl Loc COM DAT GOAL All also Abl

　　FE=Manner → GF=Dep, PT=AVP, Postpos=NUL

15

Corpus

- Balanced Corpus of Contemporary Written Japanese (BCCWJ)
  – the first available balanced and representative corpus of Modern Written Japanese
  – Copyright-free
  – Contains 143-milllion words of texts taken from:
    - Magazines, Newspapers, Government white papers, Books, Congress proceedings, Internet, and Textbooks

16

A search result of
JFN-KWIC Concordance Program

Display of parsed sentence

17

Interrelatedness of
FrameNet and Japanese FrameNet

Merge approach

- Independent resources for different languages are first built from scratch.
- Later, links that relate selected types of components cross-linguistically are added.

Expand approach

- "[A] resource for one language, which is regarded as stable at that time, is transferred to another language"

(Lonneker-Rodman 2007: 4-5)

18

3

Our Approach

Japanese FrameNet, like the Spanish and German FrameNets, adopted the expand approach

---

Differences in Frames and Constructions between FrameNet and Japanese FrameNet

---

Differences in Frames: Coverage

Coverage of existing FN semantic frames

– Can be used for **87 %** of Japanese LUs in Core Data of Book Genre of Balanced Corpus of Contemporary Written Japanese corpus (BCCWJ)

– Very few of "missing" frames are culture-specific

Japanese LUs without frame assignment

Adjective: *arai* – coarse

Conjunction: *dakara* – therefore, *sikasi* – but, *naraba* – then, *sunawati* – thus

Adjectival noun: *kooiteki* – favorable, *toozen* – naturally, *noroma* – stupid

Verb: *asobu* – play, *muku* – face, *simeru* – make up, take up, *ki o tukeru* – be careful

Adverb: *sikkari* – firmly, *tatoeba* – for example, *ippan ni* – in general

Event noun: *otukai* – errand, *taiken* – experience, *tuukoo* – crossing, *syuppan* – publication

Noun: *kami* – god, *gangu* – toy, *tan'i* – unit, *wariai* – ratio, *inu* – dog, *tatami* – straw mat, *syoozi* – sliding paper, *husuma* – sliding door, *kyookaku* – knight of the town

---

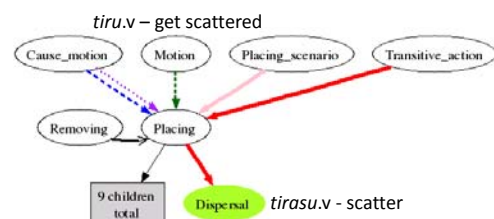## Differences in Frames: Frame Definitions

a. *sakura   no hanabira ga   **tiru** Motion*
cherry.blossom GEN petals   NOM be.scattered
'Petals of cherry blossoms get scattered.'

b. *sakura   no hanabira o   tirasu Dispersal*
cherry.blossom GEN petals   ACC scatter
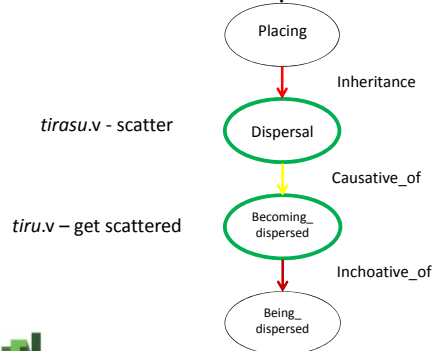'(Somebody) scatters petals of cherry blossoms.'

Frame-to-Frame Relations pertaining to `Motion` and `Dispersal` frames

Suggested Frame-to-Frame Relations for Japanese

Placing

tirasu.v - scatter

Dispersal

Inheritance

Causative_of

tiru.v – get scattered

Becoming_ dispersed

Inchoative_of

Being_ dispersed

25

Differences in Frames: Frame Definitions (Continued)

- Existing frames often assume English perspective & lexical aspects
  - Many existing frames have **transitive** perspective
    - Many **Japanese** verbs: **intransitive/inchoative** perspective
      - Transitive counterparts: often derived by suffixing a causative morpheme
  - Few cases in which current FN frames are defined from **intransitive/inchoative** *and* **transitive** perspectives
    Exception: `Becoming_detached` frame

    | `Detaching` frame | `Filling` frame |
    | `Being_detached` frame | `Fullness` frame |

26

Differences in Constructions? Types of "Multiwords" in Japanese

a. Support Verb constructions
   e.g. *mondai ni suru, ki o tukeru*
b. Complex particles
   e.g. *ni watatte, to yuu to, to itte, no toori, tokoro ni yoru to/yore ba/de wa*
c. Complex auxiliaries
   e.g. *te iru, to suru, ba ii, to yuu wake da, ni yoru*
d. Clause-level constructions
   e.g. *… yori …, sika … nai, no wa … no koto da*

=> Support Verb Constructions can be treated in the JFN lexicon but NOT others

27

Types of "Multiwords" in Japanese

a. Support Verb constructions
   mondai   ni   suru
   question DAT do
   'bring something into question'
   ki   o   tukeru
   energy ACC attach
   'pay attention to'
b. Complex particle
   kiku tokoro ni yoru   to   Hyundai wa subarasii rasii
   hear place LOC depend QUOTE   TOP superb seem
   literal. 'Depending/Based on (the) place (where I hear), Hundai seems superb.'
   = 'Judging from (what I) hear, Hyundai seems superb.'

28

Types of "Multiwords" in Japanese (continued)

c. Complex auxiliary
   koi kiri ga   numa no ue ni ori te iru
   thick fog NOM mire GEN top LOC fall
   'A thick fog has fallen over the mire.'

d. Clause-level construction
   kore (*no hoo*) *ga   are yori nagai*
   This GEN side NOM that than long
   'This is longer than that.'

29

The FrameNet Constructicon

- Goals
  - Repository of constructions
  - Not a full-fledged construction grammar
    (though it should be useful for building one)
    - "cherry-picking approach"
- Parts
  - List of constructions
  - Construction definitions
    - Description of syntactic, semantic, and pragmatic aspects of constructions and their parts
  - Annotation of sentences illustrating the constructions ("constructicographic annotation")

30

5

FrameNet Constructicographic Annotation

- Identify and annotate:
  - CONSTRUCT
    - A phrase licensed by the rules of a construction
  - Phrase type of the construct
  - CONSTRUCT ELEMENTS (CE: components of the construct) and their phrase types
  - A special CONSTRUCTION-EVOKING ELEMENT (CEE)
    - Not all constructions have such an element
    - Analogous to the target LU in FN annotation
  - Relevant features of the context

(Lee-Goldman & Rhodes 2009)

Constructional Analysis

Hearsay construction

- The *Hearsay* construction indicates that the speaker has an Attitude about the Proposition attributed to a Source.
- Construct Elements: Source, Proposition, Attitude
- [$_{Source}$ *kiku*] [$_{CEE}$ <u>*tokoro ni yoru to*</u>] [$_{Proposition}$ *Hyundai wa subarasii*][$_{Attitude}$ *rasii*] 'Judging from (what I) hear, Hyundai seems superb.'

Constructional Analyses (continued)

Resultant_state construction

- The *Resultant_state* construction describes a state after an Event pertaining to an Entity has occurred.
- Construct Elements: Entity, Event
- [$_{Entity}$ *koi kiri ga*] *numa no ue ni* [$_{Event}$ *orite*] [$_{CEE}$ <u>*iru*</u>] 'A thick fog has fallen over the mire.'

Constructional Analysis (continued)

Comparative_inequality construction

(cf. Hasegawa et al. 2010, 2012)

- The *Comparative_inequality* construction reports inequalities between two entities as arguments of a plain adjective.
- Construct Elements: Entity1, Entity2, Feature
- [$_{Entity1}$ *kore (no hoo) ga*] [$_{Entity2}$ *are*] [$_{CEE}$ <u>*yori*</u>] [$_{Feature}$ *nagai*] 'This is longer than that.'

Constructions evoking frames and appearing with Frame Elements

Hearsay construction evokes the Attribute_information frame:

- A PROPOSITION is attributed to a SPEAKER or a TEXT.
- [$_{Source}$ *kiku*] [$_{CEE}$ <u>*tokoro ni yoru to*</u>] [$_{Proposition}$ *Hyundai wa subarasii*] [$_{Attitude}$ *rasii*]
- 'Judging from (what I) hear, Hyundai seems superb.'

Constructions evoking frames and appearing with Frame Elements (Continued)

Resultant_state construction evokes the State_continue frame:

- The ENTITY remains in the specified STATE.
- [$_{Entity}$ *koi kiri ga*] *numa no ue ni* [$_{Event}$ <u>*orite*</u>] [$_{CEE}$ *iru*] 'A thick fog has fallen over the mire.'

Constructions evoking frames and appearing with Frame Elements (Continued)

Comparative_inequality construction evokes the `Comparison_inequality` frame:

- The ENTITY is compared against some STANDARD with respect to their values for some FEATURE.
- [_Entity1_ kore (*no hoo*) *ga*] [_Entity2_ *are*] [_CEE_ *yori*] [_Feature_ *nagai*]
  'This is longer than that.'

Varieties of Constructions

- Frame-bearing constructions
  - *verb-way* construction
  - *Rate* construction
  - *Reciprocal Relation* construction
- Valence-augmenting constructions
  - *Comparative –er* construction
- **Constructions without Meanings**
  - *Aux-initial* construction
- Contextually bound constructs
  - *Uniquness* construction
- Exocentric and headless Constructions
  - *Rate* construction
- Pumping Constructions
  - *Valence-modifying* constructions
- Clause-defining constructions
  - *Subject-predicate* constructions

(Fillmore et al. To Appear) 38

Purely Syntactic Constructions?

- "[w]hile construction grammars originated in the recognition of conventional pairings between specific formal patterns and the meanings they contribute to the expressions that contain them – against a contrasting view that syntactic principles should be stated independently of questions of meaning and use – there remains the question of whether all constructions should be seen as meaning-bearing."
- "the actual work of building the FrameNet Constructicon is proceeding under an assumption of the legitimacy of semantically null constructions"

(Fillmore et al. To Appear: 13)

39

Implications of Lexicon & Constructicon Building for Frame Semantics and CxG

40

How do Lexicon Building and Constructicon Building Relate to **Frame Semantics** & **Construction Grammar?**

- As a testing ground for the theories
  - Testing the methods to integrate lexical meanings and grammatical meanings into a complete account of the language-based interpretations of texts
- Refining frameworks and concepts of the theories
  - Frame-to-Frame Relations
  - "Purely Syntactic Constructions"

Summary

- Reported on difficulties encountered in developing Japanese FrameNet
- Discussed frames and constructions in English and Japanese
- Provided outlook of Japanese constructicon
- Discussed implications of lexicon and constructicon building for the sister theories
  - Frame Semantics
  - Construction Grammar

42

Conclusions

- Japanese Constructicon building is worth pursuing
  - The JFN project aims to describe relations between sentential, lexical, and constructional meanings

- Lexicon and constructicon building is all about practicing Frame Semantics and Construction Grammar

- Do it!

43

Selected References

- Baker, Collin. (2006). "Frame Semantics in Operation: The FrameNet Lexicon as an Implementation of Frame Semantics" ICCG4.
- Fillmore, Chalres J. (2006). "The Articulation of Lexicon and Constructicon" ICCG4.
- Fillmore, Charles J. (2008). "Border Conflicts: FrameNet meets construction grammar." In E. Bernal & J. DeCesaris (Eds.), *Proceedings of the XIII Euralex International Congress* (49-68). Barcelona: Institut Universitari de Lingüistica Aplicada.
- Fillmore, Charles J. and Collin F. Baker (2010) "A Frames Approach to Semantic Analysis," *The Oxford Handbook of Linguistic Analysis*, Heine, Bernd and Heiko Narrog (eds.), OUP.
- Fillmore, Charles J., Russell R. Lee-Goldman, and Russell Rhodes. (To Appear). "The FrameNet Constructicon" Boas, H.C. and Sag, I.A. (Eds.) *Sign-based Construction Grammar.*
- Hasegawa, Yoko, Russell Lee-Goldman, Kyoko Ohara, Seiko Fujii, and Charles J. Fillmore. (2010). On expressing measurement and comparison in Japanese and English. In *Contrastive Construction Grammars*, ed. Hans C. Boas, 169-200. Amsterdam: John Benjamins.
- Hasegawa, Yoko, Russell Lee-Goldman, Kyoko Hirose Ohara, Michael Ellsworth, and Charles J. Fillmore. (2012). The Frames-and-Constructions Approach to Paraphrase. ICCG7.
- Lee-Goldman , Russell & Russell Rhodes. (2009). "Corpus-based analysis and annotation of constructions" Paper presented at Frames and Constructions Conference, 31 July-2 August 2009.
- Lonneker-Rodman, Birte. (2007). "Multilinguality and FrameNet" International Computer Science Institute Technical Report.
- Ohara, Kyoko Hirose. (2008). "Lexicon, Grammar, and Multilinguality in Japanese FrameNet". LREC2008.
- Ohara, Kyoko Hirose. (2009). "Frame-based contrastive lexical semantics in Japanese FrameNet: The case of *risk* and *kakeru*". In Boas, Hans, C. (ed.), Mulitilingual FrameNets in Computational Lexicography: Methods and Applications. Mouton de Gruyter. pp. 163-182.
- Ohara, Kyoko Hirose, Hiroaki Saito, Seiko Fujii, and Hiroaki Sato. (2011). "Japanese FrameNet: Building a lexical and constructional resource based on BCCWJ and Semantic Frames." (In Japanese).
- Ohara, Kyoko Hirose (2012). "Semantic Annotations in Japanese FrameNet: Comparing Frames in Japanese and English ". LREC2012.
- Ruppenhofer, Joseph, et al. (2010). "FrameNet II: Extended Theory and Practice."

44

URLs

- Japanese FrameNet http://jfn.st.hc.keio.ac.jp/

- JFN data on FrameSQL
  http://framenet2.icsi.berkeley.edu/frameSQL/jfn23/notes/index.html

- Japanese FrameNet on YouTube

  http://www.youtube.com/watch?v=kfqR9aUcp1c

45

46