Frame-based analysis of polysemous words in Japanese FrameNet

JAPANESE FRAMENET

Kyoko Hirose Ohara

Keio University

ohara@hc.st.keio.ac.jp

Outline

1. Purpose of the talk

2. What is Japanese FrameNet

3. How is Frame-based Japanese FrameNet different from other linguistic resources

4. Summary

Purpose of the talk

- Preliminary report on an on-going study

  ➢ Comparing lexical entries of Japanese FrameNet (JFN) with those in other linguistic resources

  ➢ To Test the hypothesis "Frame-based analysis is suited for describing related word senses"
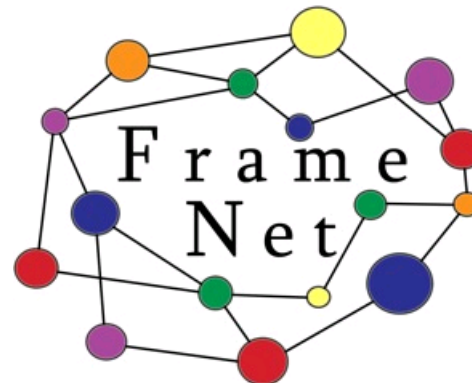
What is Japanese FrameNet

Japanese FrameNet (JFN)

- Creating a **prototype of an on-line Japanese linguistic resource** following FrameNet methodology and practice

- Compatibility with FrameNet

  - JFN databases and annotation tool

  - JFN frames: imported from FN (the Expand approach)

  - Annotation methods

- Lexicon building > Full Text Annotation > Constructicon Building

JFN Annotation Tool

Lexical Unit

Postposition Layer
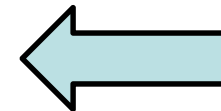
Valence Patterns in JFN

Consist of:

- Frame Elements (FEs)

- Grammatical Functions (GFs: Subj, Obj, …)

- Phrase Types (PTs)

- *Postpositions* (i.e., "surface" case)

JFN Output: Lexical Entry Report

8

JFN Output: FrameSQL

Sato (2012)

JFN Output:
Full Text Annotation

全文テキストアノテーション

[PB56_00002.txt]

Current Status

- Lexicographic annotation
  - # of LU's: 3,000 -
  - # of Annotated Sentences: 60,000 -
- Full Text Annotation
  - # of Annotated Sentences: 1,000 -

- Constructicon Building
- JFN Web Annotation Tool (JFNWAT)

JFN Web Annotation Tool (JFNWAT)

Being developed by Interns from Ecole Centrale de Lyon

Current Grants
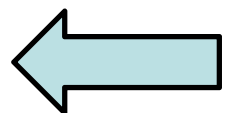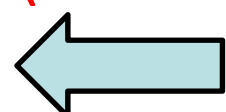
- Constructicon Building
  - Grant-in-Aid for Scientific Research
    - 2012-2015
- Full Text Annotation
  - Grant from the National Institute of Japanese Language and Linguistics
    - 2012-
- JFN Data Release
  - Grant-in-Aid for Scientific Research
    - 2013-2017

How is Frame-based
Japanese FrameNet
different from
other linguistic resources

"Predicate frame lexicons" in Japanese

- IPAL (Information-Technology Promotion Agency , 1987)

- EDA (Japan Electronic Dictionary Research Institute, 1995)

- GoiTaikei (Ikehara et al., 1997)

- Japanese FrameNet (Ohara et al., 2003)

- Kyoto University Case Frames (Kawahara & Kurohashi, 2006)

- Verb Argument Structure Thesaurus (VAST) (Takeuchi et al., 2010)

15

Comparing VAST & JFN entries

- ## VAST
  - "splitter" approach rather than "lumper" approach, just like JFN?!

- ## Method
  - Collected 109 lemmas in JFN having multiple LU's
  - Compared VAST & JFN entries

- ## Hypothesis
  - Frame-based approach of Berkeley FrameNet and Japanese FrameNet (inter alia) is better suited to describing related word senses, due to
    - Fine-grained Frame Elements, relativized to frames
    - Rich Valency annotations

Case Study 1
VAST entries for *omou* 'think, ponder'

(1) Takeuchi et al. 2010

   a.   <**Experiencer**> *ga*   <**Target**> *o* <**Descriptive**> *omou*

                   NOM   ACC   think

   b. <**Agent**> *ga*   <**Target**> *o*   *omou*

              NOM   ACC imagine

   c. <**Agent**> *ga*   <**Target**> *o*   *omou*

              NOM   ACC ponder

JFN frames for *omou*.v

(2) `Cogitation` frame

    A **COGNIZER**, thinks about a **TOPIC** over a period of time.

(3) `Coming_to_beleive` frame

    A **COGNIZER**, comes to believe something (the **CONTENT**), sometimes after a process of reasoning.

(4) `Deciding` frame

    A **COGNIZER** makes a **Decision**, which may be about an entity or a course of action.

(5) `Expectation` frame

    A **COGNIZER**, believes that some **Phenomenon** will take place in the future.

(6) `Opinion` frame

    A **COGNIZER** holds a particular **Opinion**, which may be portrayed as being about a particular **Topic**.

Matching VAST and JFN entries for *omou*

a. <**Experiencer**> *ga*   <**Target**> *o* <**Descriptive**> *omou*

           NOM   ACC   think

   ⇒ Opinion frame

b. <**Agent**> *ga*   <**Target**> *o*   *omou*

         NOM   ACC   imagine

   ⇒ Cogitation frame

c. <**Agent**> *ga*   <**Target**> *o*   *omou*

         NOM   ACC   ponder

   ⇒ Cogitation frame

Comparing VAST and JFN entries for *omou*

- Among the 5 LU's for *omou*.v in JFN, `Cogitation`.*omou*.v is the only LU with postpositions **_ga_ (NOM) + o (ACC)** alone. The other uses take the quotative marker *to.*

- *watasi   -tati **ga**   aisuru hito-tati no   koto **o** <u>omou</u> toki …*

  1st person PL.   NOM love   person Pl. GEN thing ACC   when

  "When we think about the people whom we love, …"

  VAST does not reflect the fact above

Case Study 2
VAST and JFN entries for *kangaeru* 'think'

(7)   Entry for *kangaeru.v* 'think' in VAST

<Experiencer> *ga*   <Content> *o*   <u>*kangaeru*</u>

NOM   ACC

e.g.   [<Content> *kono syoosetu no teema   o*]   <u>*kangaeru*</u>

this novel   GEN theme ACC

'(I) think about the theme of this novel.'

(8)   Cogitation frame in JFN

A **COGNIZER** thinks about a **TOPIC** over a period of time.

(9)   Cogitation.*kangaeru*.v in JFN

[**COGNIZER** External NP *ga*] [**TOPIC** Dependent **NP/S**   *o/ni.tuite*] <u>*kangaeru*</u>

NOM   ACC/about

JFN annotation for *kangaeru*.v

(10) *Zen.sekai ni.watatte* [TOPIC *kono mondai ni.tuite*] *kangaeru*
all.world throughout   this problem about

beki da

should COP

'Throughout the whole world, (we) should consider this problem.'

(11) [TOPIC *dare ga   daihyoo   to.site husawasii ka*] *kangaeru*

who NOM representative as   appropriate Q

'think who would be appropriate as the representative'

⇒ JFN records the fact that `Cogitation`.*kangaeru*.v allows the
TOPIC to accompany the postposition *ni.tuite* and to be realized as a
sentential complement (PT=S), but *not* VAST.

22

Case study 3
Describing relations between senses

- *hanako wa **taroo ni***   *tyokoreeto o*   *ageta*

     TOP   DAT chocolate   ACC gave

     'Hanako gave chocolate to Taro.'

- *hanako wa **taroo ni***   *tyokoreeto o*   *moratta*

     TOP   DAT   ACC received

     'Hanako received chocolate from Taro.'

JFN entries for *ageru* 'give' & *morau* 'receive'

– Giving frame

- A Donor transfers a Theme from a Donor to a Recipient.

- *hanako wa **taroo ni** tyokoreeto o ageta*

     TOP   DAT chocolate   ACC gave

  'Hanako gave chocolate to Taro.'

– Receiving frame

- A Recipient comes into possession of the Theme as a result of the joint action of the Donor and the Recipient.

- *hanako wa **taroo ni** tyokoreeto o moratta*

     TOP   DAT   ACC received

  'Hanako received chocolate from Taro.'

24

JFN valence patterns for
ageru 'give' & *morau* 'receive'

- Giving frame

  - *ageru*.v

    {[Donor .EXT.NP.*ga*]

    [Recipient .Dep.NP.***ni***]

    [Theme .Obj.NP.*o*]}

- Receiving frame

  - *morau*.v

    {[Recipient .EXT.NP.*ga*]

    [Donor .Dep.NP.***ni***]

    [Theme .Obj.NP.*o*]}

Frame-to-Frame relations
between `Giving` & `Receiving` frames

Perspective_on   Transfer frame   Perspective_on

Giving frame   Receiving frame
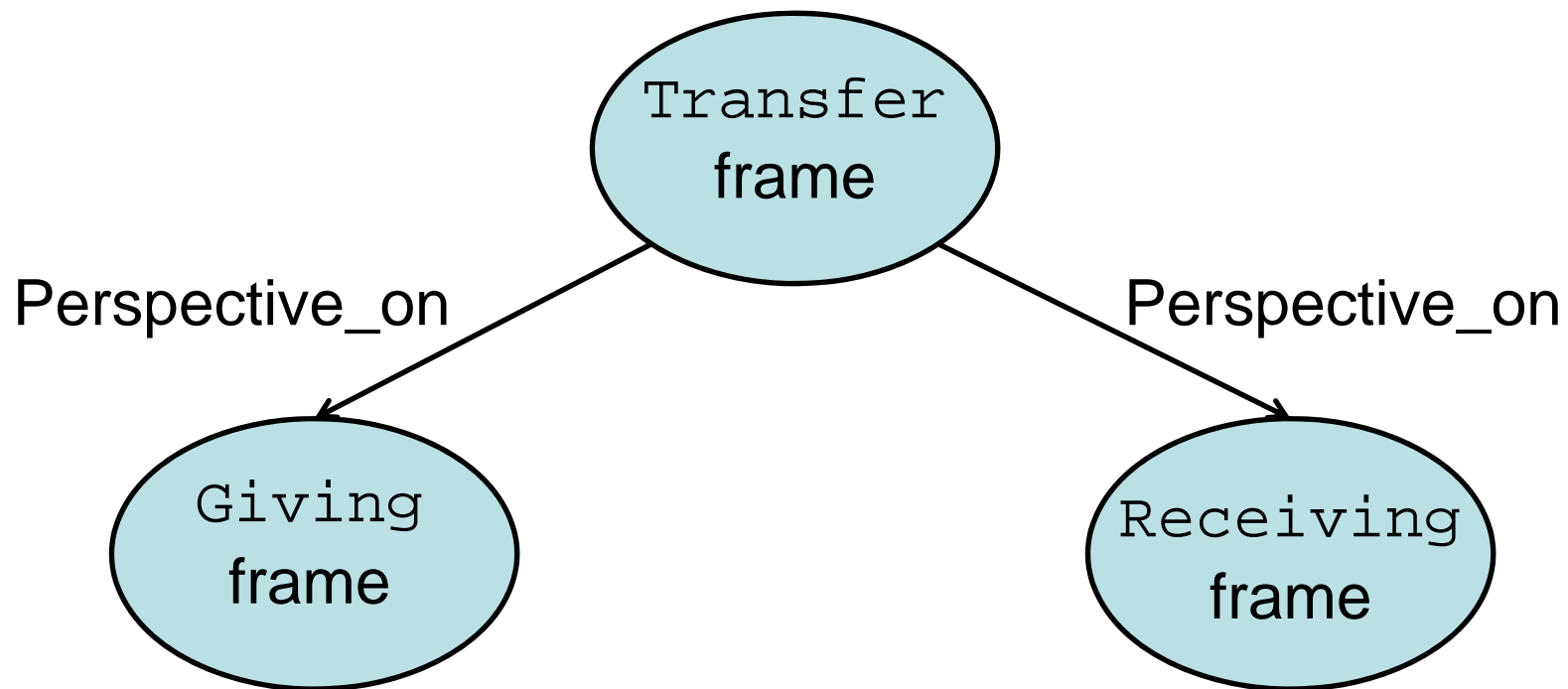
⇒ VAST cannot describe relations between senses

Summary

- Frame-based approach, i.e., rich valency annotation, especially information about **PT *and Postposition* (i.e., "surface" case)**, is better suited for describing fine-grained senses of polysemous words.

- Frame-to-Frame relations help to describe relations between senses.

- Future directions
    - Compare annotations (cf. de Melo et al. 2012)
    - Use JFN in applications such as Word Sense Disambiguation, Paraphrasing, and Machine Translation (cf. Hasegawa et al. 2011, 2012)

27

Thank You!

- Japanese FrameNet is supported in part by
  - Grant-in-Aid for Scientific Research (C) #24520437
  - Grant from the National Institute of Japanese Language and Linguistics

Selected References

- De Melo, Gerald, and Collin F. Baker. (2012). Empirical Comparisons of MASC Word Sense Annotations: a First Look. LREC2012.

- 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.

- Hasegawa, Yoko, Russell Lee-Goldman, Albert Kong, and Kimi Akita. (2011). FrameNet as a resource for paraphrase research. *Constructions and Frames* vol. 3, no. 1, 104-127.

- Hasegawa, Yoko, Russell Lee-Goldman, Kyoko Hirose Ohara, Michael Ellsworth, and Charles J. Fillmore. (2012). The Frames-and-Constructions Approach to Paraphrase. ICCG7.

- Hasegawa, Yoko, Russell Lee-Goldman and Charles J. Fillmore (Forthcoming). Universality of frames: Evidence from English-to-Japanese translation. In Jan-Ola Ostman and Mirjam Fried (eds.), *The Socio-Pragmatics of Frame Semantics and Construction Grammar*. Amsterdam: John Benjamins.

- Ikehara, Satoru, Masahiro Miyazaki, Satoshi Shirai, Akio Yokoo, Hiromi Nakaiwa, Kentaro Ogura, Yoshifumi Ooyama, and Yoshihiko Hayashi. (1997). *Goi-Taikei — A Japanese Lexicon*, Iwanami Shoten, Tokyo.

- Information-Technology Promotion Agency. (1987). Japanese Verbs: A Guide to the IPA Lexicon of Basic Japanese Verbs.

- Japan Electronic Dictionary Research Institute. (1995). EDR Electronic Dictionary Technical Guide (2nd edition), Technical report, Japan Electronic Dictionary Research Institute, Ltd., Tokyo.

- Kawahara, Daisuke, and Sadao Kurohashi. (2006). Case frame compilation from the web using high-performance computing, in *Proceedings of the 5th International Conference on Language Resources and Evaluation (LREC 2006)*, pp. 1344–1347, URL http://nlp.ist.i. kyoto-u.ac.jp/member/kawahara/paper/LREC2006.pdf.

- Lonneker-Rodman, Birte. (2007). "Multilinguality and FrameNet" International Computer Science Institute Technical Report.

- Ohara, Kyoko, Seiko Fujii, Hiroaki Saito, Shun Ishizaki, Toshio Ohori, and Ryoko Suzuki. (2003). The Japanese FrameNet Project: A Preliminary Report, in *Proceedings of Pacific Association for Computational Linguistics (PACLING'03)*, pp. 249–254.

- Ohara, Kyoko Hirose (2012). "Semantic Annotations in Japanese FrameNet: Comparing Frames in Japanese and English ". LREC2012.

- Takeuchi, Koichi, Kentaro Inui, Nao Takeuchi, and Atsushi Fujita. (2010). A Thesaurus of Predicate-Argument Structure for Japanese Verbs to Deal with Granularity of Verb Meanings, in *Proceedings of the 8th Workshop on Asian Language Resources*, pp. 1–8.

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