

# Investigating Japanese FrameNet Data with FrameSQL



Kyoko Hirose Ohara & Hiroaki Sato  
Keio University & Senshu University, Japan  
ohara@hc.cc.keio.ac.jp, thc0408@gmail.com

5<sup>th</sup> September, 2010  
Charles University, Prague, Czech Republic

## Outline

1. Introduction
  2. FrameNet (BFN) and Japanese FrameNet (JFN)
  3. FrameSQL
  4. Analyses of *surprise.v* & *odorokasu.v*
  5. Conclusions
  6. Future Work
- References

2

## 1. Introduction

- Existing lexical resources
  - Still a widespread assumption that a small number of universally-defined semantic roles are sufficient and more preferable than frame elements (FE's) defined in terms of specific semantic frames
    - E.g. “Verb Argument Structure Thesaurus” (VAST) (Takeuchi et al. 2010)
    - But cf. Matsubayashi et al. (2010)

3

## 1.1. Claims

- JFN is more suitable to analyze actual uses of words in corpora than other existing lexical resources, because in JFN annotations,
  - 1) not only core frame elements but also non-core frame elements are used;
  - 2) the information about which frame element is paired with which grammatical functions (GF's)/phrase types (PT's) is recorded as valence;
  - 3) frame elements are linked to each other by frame-to-frame relations.

4

## 2. FrameNet and Japanese FrameNet 2.1. FrameNet

- “The general purposes of the [Berkeley FrameNet] project are both to provide reliable descriptions of the syntactic and semantic combinatorial properties of each word in the lexicon, and to assemble information about alternative ways of expressing concepts in the same domain.” --- Fillmore & Baker (2010:321)

5

## 2.2. Japanese FrameNet

- Creating a **prototype of an online Japanese lexical resource** in the FrameNet style:
  - By describing the **sense of each lexical unit** with respect to the **semantic frame** it evokes;
  - By **annotating corpus examples** of each word with frame-semantic tags (frame elements).
- Research questions
  - To what extent is the Frame-semantic approach suitable for analyzing the Japanese lexicon?
  - To what extent are the existing English-based semantic frames applicable to characterizing Japanese lexical units?
- Currently annotating Balanced Corpus of Contemporary Written Japanese (BCCWJ)

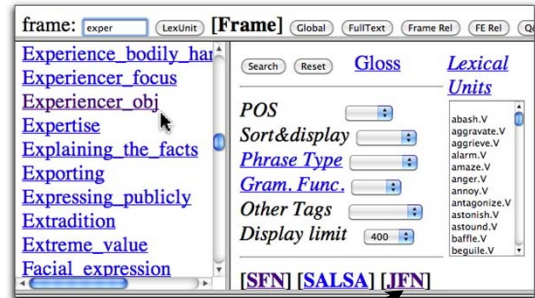
6

### 3. FrameSQL

- a web-based application created originally for the Berkeley FN.
- puts together JFN and BFN lexical databases, and the user can access them seamlessly, as if they were a unified database.
- useful for comparing related frames and lexical units.

7

### 3.1. BFN search menu



- [hyper link to JFN](#)

8

### 3.2. JFN search results

Link to Verb Argument Structure Thesaurus (VAST)

No.	frame	lemma	pos	def.	動詞項構造シソーラス (VAST)
1	Experiencer_obj	驚かす	V	nil	驚かす

Display Modes: ALL FE KDG valence

Num FE/LUset (sort = KDG; Experiencer\_obj, 驚かす, V,)

01 [Circumstances/繰り返し巻き返し内部から起こる改革運動エネルギーを秘めていたのに]+驚かさ+[(Experiencer)/

01 [Circumstances/改めて]+[Stimulus/マスコミのあのまづき]+[Experiencer]

### 4. Analyses of

#### surprise.v & odorokasu.v

The Experiencer\_obj frame in BFN & JFN:

Some phenomenon (the **STIMULUS**) provokes a particular emotion in an **EXPERIENCER**.

Core FE's:

- **EXPERIENCER**: The **EXPERIENCER** reacts emotionally or psychologically to the **STIMULUS**.
- **STIMULUS**: The **STIMULUS** is the event or entity which brings about the emotional or psychological state of the **EXPERIENCER**.

Non-core FE's:

- **MEANS**: The **MEANS** by which the **STIMULUS** affects the **EXPERIENCER**.
- **TIME**: The **TIME** when the **EXPERIENCER** has an emotion as caused by the **STIMULUS**.

10

### 4.1. FN: Phrase Types of **STIMULUS X**

1. NP: **X** surprised Y
2. PP: Y was surprised by **X** (Passive)
3. Sfin: It surprised Y [**STIMULUS** that S] (Extraposition)
4. VP<sub>to</sub>: It surprised Y [**STIMULUS** to VP] (Extraposition)
5. when: ... it always surprises me [**STIMULUS** when people turn out to be such bad listeners]
6. if: ... it wouldn't surprise me in the least [**STIMULUS** if, if you analyzed this dream you discovered it had nothing to do with being at school and nothing to do with taking exams]

### 4.2. FN: when clauses

with *devastate.v* in the Experiencer\_obj frame

1. [**EXPERIENCER** A family I knew] were devastated [**STIMULUS**+ **TIME** **when the thirty-five-year-old mother died after a short illness**].

12

### 4.3. Frame-to-Frame Relations

- “every FE is defined in terms of specific frames”
- Problem: too many FEs
  - “At the moment (2008.10.1) there are slightly more than **9,000** distinct types of Frame Element in FrameNet.”
  - FrameNet II Forum, [http://framenet.icsi.berkeley.edu/index.php?option=com\\_mamboboard&Itemid=141&func=view&catid=5&id=253](http://framenet.icsi.berkeley.edu/index.php?option=com_mamboboard&Itemid=141&func=view&catid=5&id=253)
- Solution: Frame-to-Frame Relations
  - “But don't panic! They are related to each other via [several kinds of] relations, and in most cases can be recognized as a subtype of a small set of FEs that resemble thematic roles.”

13

### 4.4. Relations of **STIMULUS**

Emotions.Expressor >>	<< Contrition.Action
Emotions.Manner >>	<< Desiring.Event >> (Using)
Emotions.Parameter >>	<< Emotion_directed.Stimulus
Emotions.Reason >>	<< Emotions_by_stimulus.Stimulus >>
Emotions.State >>	<< Experiencer_obj.Stimulus
<b>Emotions.Stimulus &gt;&gt;</b>	<< Forgiveness.Offense >> (Using)
Emotions.Topic >>	<< Judgment.Evaluatee >> (Using)
<<	<< Predicament.Situation
Emotions_by_possibility.Circ	<< Stimulus_focus.Stimulus >> (Persp

- Emotions. **STIMULUS** shares some properties with other related FEs.

14

### 4.5. [**STIMULUS** when ...] in the Emotions\_by\_stimulus frame

1. LUs: glad, joyful
2. Sfin: X was glad [**STIMULUS** that S]
3. VPto: X was glad [**STIMULUS** to VP]
4. when: We were as glad as anything [**STIMULUS**+ TIME **when we saw this rain coming down**].

15

### 4.6. seeing with **STIMULUS**

1. [**STIMULUS** Seeing Sadako's dance] (*Sadako no odori o mi-TE*)  
GEN dance ACC see-TE  
Toyama was much surprised.
  2. [**MEANS** Starting the game] (*geemu o hajime-TE*)  
game ACC start-TE  
Deumi was surprised.
- Question: Is there any data to support this distinction?
  - Answer: There are currently only three LUs in the Experiencer\_obj frame in JFN, but Frame-to-Frame relations give us more data in related frames.

16

### 4.7. seeing with **CONTENT**

1. Experiencer\_obj <=>>  
Experiencer\_focus  
• *odorokasu* (surprise.v) <=>>  
*odoroku* (be\_surprised.adj)
2. Experiencer\_obj. STIMULUS <=>>  
Experiencer\_focus . CONTENT
3. [**CONTENT** Seeing his hands] (*Jibun no te o mi-TE*)  
Mr. Maeda was surprised.  
• JFN tries to annotate the data consistently.

17

### 4.8. Annotating dictionary example phrases

- (1) *odorokasu.v* 'to surprise' in *Daijirin: The Second Edition* (2006)
- a. *seken o odorokaseta ziken*  
public ACC surprised incident  
'the incident which surprised the public'
  - b. *zimoku o odorokasu*  
many\_people's\_attention ACC surprise  
'to surprise people'

18

## 4.9. Relevant entries in VAST and JFN

(2) VAST entry for *odorokasu* (Takeuchi et al. 2008)

a. <Agent> *ga* <Person> *o* *odorokasu*  
NOM ACC surprise

b. <Causer> *ga* <Person> *o* *odorokasu*

(3) The Experiencer\_obj frame in JFN

Some phenomenon (the **STIMULUS**) provokes a particular emotion in an **EXPERIENCER**.

19

## 4.10. Annotations of (1) in VAST and JFN

(2') VAST annotations

a. [< Person > *seken o*] *odorokaseta* [< Causer > *ziken*]  
public ACC surprised incident  
'the incident which surprised the public'

b. [< Person > *zimoku o*] *odorokasu*  
many\_people's\_attention ACC surprise  
'to surprise people'

(3') JFN annotations

a. [EXPERIENCER *seken o*] *odorokaseta* [STIMULUS *ziken*]

b. [EXPERIENCER *zimoku o*] *odorokasu*

20

## 4.11. Summary:

### Treatment of 'peripheral' phrases

- Our analysis of *odorokasu* (to surprise) shows that JFN assigns FEs to adjunct phrases, which are often disregarded as 'peripheral' in VAST.
- Many sentences in the corpus contain adjunct phrases, and JFN uses the framework of Frame Semantics to describe them properly, just as BFN does.
  - JFN: [STIMULUS **Seeing Sadako's dance**] (*Sadako no odori o mite*), Tooyama was much surprised.
  - BFN: ... it always surprises me [STIMULUS **when people turn out to be such bad listeners**]

21

## 4.12. Annotating Corpus Sentences

(4) Sentence from the BCCWJ corpus

*Sadako ga dansu o suru siin o soozosita koto mo*  
NOM dance ACC do scene ACC imagined thing PART  
*nakatta tame, Sadako no odori o mite,*  
did.not.exist SUB GEN dance ACC see-TE  
*Tooyama wa kanari odorokasareta.*  
TOP much be.surprised

'Since (he) had not imagined a scene in which Sadako performs a dance, seeing her dance, Toyama was much surprised.'

22

## 4.13. Annotations of (4) in VAST and JFN

(5) VAST annotation

*Sadako ga dansu o suru siin o soozosita koto mo*  
*nakatta tame,*  
*Sadako no odori o mite,*  
[< Person > *Tooyama wa*]  
*kanari odorokasareta.*

(6) JFN annotation

[EXPLANATION *Sadako ga dansu o suru siin o soozosita koto mo*  
*nakatta tame*]  
[STIMULUS *Sadako no odori o mite*],  
[EXPERIENCER *Tooyama wa*]  
[DEGREE *kanari*] *odorokasareta.*

23

## 5. Conclusions

- FrameSQL highlights interesting and important valence patterns hidden in corpus data, showing that the FrameNet/Frame Semantics approach is suitable for analyzing Japanese texts.

24

## 6. Future Work

- Constr. Name: *When- STIMULUS* Construction
  - Form: [**STIMULUS** *when* S] ... [<Emotion-related frame> LU]
  - Meaning: **EXPERIENCER** experiences an emotion when **S STIMULUS**.
- Constr. Name: **STIMULUS –TE** Construction
  - Form: [**STIMULUS** ... VP-TE] ... [<Emotion-related frame> LU]
  - Meaning: **EXPERIENCER** experiences an emotion when VP **STIMULUS**.

25

## References

- 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 Lingüística 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.
- Matsubayashi, Yuichiroh, Naoagki Okazaki, and Jun'chi Tsujii. (2010). "Generalization of Semantic Roles in Automatic Semantic Role Labeling." (In Japanese) *Journal of Natural Language Processing*. Vol. 17, No. 4.
- Ruppenhofer, Joseph, et al. (2010). "FrameNet II: Extended Theory and Practice."
- Sato, Hiroaki. (2010). "Viewing the Japanese FrameNet Data with FrameSQL." (In Japanese). In *Proceedings of the Workshop on the Balanced Corpus of Contemporary Written Japanese*. 143-146.
- Koichi Takeuchi, Kentaro Inui, Nao Takeuchi and Atsushi Fujita (2010). A Thesaurus of Predicate-Argument Structure for Japanese Verbs to Deal with Granularity of Verb Meanings, The 8th Workshop on Asian Language Resources, pages 1-8, Beijing.

26

## URLs

- Balanced Corpus of Contemporary Written Japanese (BCCWJ) website  
[http://www.kokken.go.jp/en/research\\_projects/kotonoha/bccwj/](http://www.kokken.go.jp/en/research_projects/kotonoha/bccwj/)
- FrameNet website  
<http://framenet.icsi.berkeley.edu/>
- FrameSQL website  
<http://sato.fm.senshu-u.ac.jp/jfn23/notes/index.html>
- Japanese FrameNet website  
<http://jfn.st.hc.keio.ac.jp/>
- Japanese FrameNet on YouTube  
<http://www.youtube.com/watch?v=kfqR9aUcp1c>

27