

Frame-based analysis of polysemous words in Japanese FrameNet



JAPANESE FRAMENET

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



19th April, 2013
International FrameNet Workshop 2013
The Shattuck Plaza Hotel, Berkeley

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”

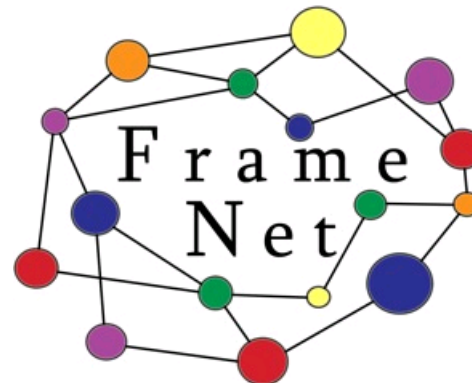


What is Japanese FrameNet



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 > Construction Building



JFN Annotation Tool

Japanese FrameNet Desktop

Main Action Window

Arrest

Arriving

- Goal <G>
- Theme <H>
- Manner <M>
- Means <F12>
- Mode_of_transportation
- Path <F5>
- Source <S>
- Time <F2>
- Cotheme <K>
- Depictive <D>
- Goal_conditions <F8>

たどりつく.v

つく.v

入る.v

来る.v

着く.v

Lemma(V)

から+着く-(2) [1/1]

が+着く-(2) [2/2]

で+着く-(2) [1/1]

と+着く-(2) [1/1]

なかなか+着く-(2) [1/1]

に+着く-(2) [9/75]

の+着く-(1) [1/1]

は+着く-(2) [1/1]

ば+着く-(2) [1/1]

へ+着く-(1) [5/5]

「いまから一時間

「するとあなたは

「さあ、話の模様

それからまもなくも

それでは夕がたま

へ+着く-(2) [6/6]

東京から約290

旅も順調に進み、

昨年8月15日夜

信楽高原鉄道に乗

新橋駅 新橋駅

Sentence Editor: 昨年 8 月 1 5 日 夜、友人 2 (12066)

Layer

FE 昨年 8 月 1 5 日 夜、友人 2 人と東京をたち、翌朝 6 時に奈良駅へ着いた。

GF

PT

Other

Verb

Sent

Case

Lexical Unit

Postposition Layer

FE GF PT Other Verb Sent Case

Goal <G>

Theme <H> DNI

Manner <M>

Means <F12>

Mode_of_transportation ...

Path <F5>

Source <S>

Time <F2>

Cotheme <K>

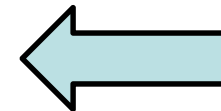
Depictive <D>

Goal_conditions <F8>

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

Arriving
たどりつく.v
つく.v
入る.v
来る.v
着く.v
至る.v

Arson
Artifact
Assessing
Assistance
Atonement
Attaching
Attack
Attempt
Attempt suasion
Attention
Attention getting
Attributes
Avoiding
Awareness
Bail setting
Bearing arms
Become silent
Becoming
Becoming aware
Behind the scenes
Being attached
Being born
Being dry
Being employed

Valence Patterns:

These frame elements occur in the following syntactic patterns:

Number Annotated	Patterns			
(1) TOTAL	Depictive	Goal	Theme	
(1)	AVP Dep	NP Dep	NP Ext	
	--	--	が	
(1) TOTAL	Depictive	Theme	Time	
(1)	Sub Dep	NP Ext	AVP Dep	
	--	は	--	
(1) TOTAL	Goal			

ColorsOff

◆そして彼はあごひげなどで変装して、モーティマー医師をここまで尾行してみたり、**ペンリー卿の着**たときは駅へも行ったし、ノーサンバランド・ホテルへ尾行したりして出歩いている間、細君はホテルの居間に監禁しておいたのだ。**DNI**

◆やっと**着**いたか。**DNI DNI**

◆相互タクシー（大阪市城東区）でも、依頼場所に到着した際に「車が着きました」という連絡が客にできず、客から「**車は着**いたのか」といった苦情が数件寄せられたという。**DNI**

◆美智子がタクシーに飛び乗ってくれたんですが、なかなか**着**かないから僕は覚悟を決めた。**DNI**

DNI

◆**携帯のメールが着**いたら、相手に自動連絡するような新手のサービスだろう。**DNI**

◆**ひつぎが着**いてから、骨壺（つぼ）を持って遺族が退場するまで通常約1時間半と見込まれて

Postposition Layer

JFN Output: FrameSQL

Sato (2012)

The screenshot shows the JFN FrameSQL interface. On the left is a list of frame names: [Commerce_pay](#), [Contingency](#), [Departing](#), [Desirability](#), [Existence](#), [Expectation](#), [Experiencer_focus](#), [Experiencer_obj](#), [Explaining_the_facts](#), [Fluidic_motion](#), [Grant_permission](#), [Motion](#), and [Motion_directional](#). The main panel has a search area with buttons for 'Search', 'Reset', and 'Gloss'. Below these are filters for 'POS', 'Sort&display', 'Phrase Type', 'Gram. Func.', 'Other Tags', and 'Display limit' (set to 9999). A red circle highlights the '[BFN]' button. To the right are two scrollable lists: 'Lexical Units' containing Japanese verbs like 'あきる.V', 'あきれる.V', 'おそれる.V', 'おどろく.V', 'ひるむ.V', '悔る.V', '喜ぶ.V', '怯える.V', '恐れる.V', '恥じる.V', '悔いる.V', and '愛す.V'; and 'Frame Elements' containing roles like 'Beneficiary', 'Circumstances', 'Content', 'Degree', 'Event', 'Event+Expressor+State+Topic', 'Experiencer', 'Experiencer+Content', 'Explanation', 'Expressor', 'Manner', and 'Parameter'.

Experiencer_focus

Definition:

The words in this frame describe an **Experiencer**'s emotions with respect to some **Content**. A **Explanation** for the emotion in current state of affairs, quite often it refers to a general situation which causes the emotion.

- **My** **ENJOYMENT** **of the movie** was considerably impaired by the seven-foot guy sitting in front of me. [Yahoo!Japan¹]
- **Smithers** **takes** **great** **PLEASURE** **in collecting matchboxes**. [Yahoo!Japan¹ 翻訳 **EJ**]

JFN Output: Full Text Annotation

全文テキストアノテーション

[PB56_00002.txt]

6. パイロット People,by,vocation 最たる 仕事 Being,obligated は 資格 Documents 維持 Activity,ongoing 「その2」 次 Relative,time に 操縦 Operate,vehicle 関係 Relation の ライセンス Documents だが、これも 管制 Control 機関 Organization と 文信 Communication する 航空 Operate,vehicle 級 Rank 無線 Communication,means 通信 People,by,vocation 士の ライセンス Documents と、操縦 Operate,vehicle そのものの ライセンス Documents とに 分かれる Becoming,separated
2. ここで 操縦 Operate,vehicle そのものの ライセンス Documents に 話 Communication を 絞つて Place,weight,as て 進める Activity,ongoing と、パイロット People,by,vocation は 訓練 Education,teaching Education,teaching 生となつ Becoming たその 日 Calendaric,unit から、約 Relational,quantity 三十 時間 Measure,duration の 飛行 Operate,vehicle 時間 Measure,duration でようやくノロ(単独。) 飛行 Operate,vehicle が できる Capability ようになる。
3. その後、自家用 操縦 People,by,vocation 士 免許 Documents 事業 Businesses 用 Purpose 操縦 People,by,vocation 士 免許 Documents 一切 Completeness 外 Part,inner,outer を 見る Perception,active ことなく 計器 Gismo のみで 飛ぶ Operate,vehicle ことができる Capability 計器 Gismo 飛行 Operate,vehicle 証明 Documents 双発以上の エンジン Vehicle,subpart を 持つ Possession 飛行機を 操縦 Operate,vehicle できる Capability 陸上 Operate,vehicle 多発機などの 基本 Being,necessary 的な 免許 Documents を 約 Relational,quantity 四 年間 Measure,duration 飛行 Operate,vehicle 時間 Measure,duration にして 約 Relational,quantity 三百数十 時間 Measure,duration をかけて 取得 Getting する。
4. ここで 言う Communication 「陸上 Vehicle 多発機」とは、海上 Relational,natural,features や 湖国 Relational,natural,features で 離着陸 Motion できる Capability 水上 Vehicle 機に対して、通常 Typicality の 滑走 Motion,Locate,by,use 路 Roadways を 使用 Using して 離着陸 Motion する飛行機を 指す Referring,by,name 用語 Simple,name と 理解 Grasping していただきたい。
5. しかし 航空 Operate,vehicle 法 Law では、これら自家用 操縦 People,by,vocation 士、事業 Businesses 用 Purpose 操縦 People,by,vocation 士、計器 Gismo 飛行 Operate,vehicle 証明 Documents 陸上多発の ライセンス Documents だけでは、まだ 最大 Extreme,value 離陸 Motion 重量 Dimension が五千七百キログラム(五、七トン。)未満の飛行機しか 操縦 Operate,vehicle でき Capability ない。
6. 「最大 Extreme,value 離陸 Motion 重量 Dimension が五千七百キログラム」の飛行機とは…？ それは具体的に、十数人乗りの飛行機を イメージ Awareness していただくといいたいだろう。
7. 一方、定期 Frequency 航空 Motion で 使用 Using している 機材 Supply は、数十名乗りの飛行機から 最大 Extreme,value 五百六十八席の ジャンボ Vehicle ジェット(最大 Extreme,value 離陸 Motion 重量 Dimension は三十八万キログラム/三百八十 トン 前後 Relational,quantity)まであり、最大 Extreme,value 離陸 Motion 重量 Dimension が五、七トン以上の大きさの飛行機を 操縦 Operate,vehicle するには、これらの ライセンス Documents ではまだまだ 不足 Sufficiency で、それらの飛行機に独自の ライセンス Documents を 取得 Getting しなければならない。
8. 最大 Extreme,value 離陸 Motion 重量 Dimension を基準に個々の飛行機に個々の ライセンス Documents が 要求 Being,necessary される 理由 Reason は、飛行機によって、それぞれの 設計 Config,us,with 思想や 飛行 Motion 特性が 大きく Degree 異なる Similarity ことによる。

その後、自家用操縦士免許、事業用操縦士免許、**一切の航空機に共通のライセンス**計器のみで**飛ぶ**ことができる計器飛行証明、双発以上のエンジンを持つ飛行機を操縦できる陸上多発機などの基本的な免許を約四年間、飛行時間にして約三百数十時間をかけて取得する。**陸上多発機**



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)

Main ▾ Action ▾ Window ▾ Logged as : urd, auth : [read] ▾

Frame Tree

- Apply_heat
- Appointing
- Architectural_part
- Arraignment
- Arranging
- Arrest
- ▾ Arriving
 - Goal
 - Theme
 - Manner
 - Means
 - Mode_of_transportation
 - Path
 - Source
 - Time
 - + Cotheme
 - + Depictive
 - + Goal_conditions
 - ① 来る.v
 - ① 至る.v
 - ① 入る.v
 - ① つく.v
 - ① 着く.v
 - ① たどりつく.v
 - ① くる.v
 - ① 帰る.v
 - ① なる.v
 - ① 近づく.v

Report Frame Arriving

Arriving

Definition

An object Theme moves in the direction of a Goal. The Goal may be expressed or it may be understood from context, but its is always implied by the verb itself.

Semantic Types

FES

Core	
Theme	Theme is the object that moves. It may be an entity that moves under its own power, but it need not be. <i>The officer approached the house.</i> <i>I ducked as the baseball approached my head.</i> <i>十二時前になって新たな客が入ってきた。</i>
Goal	Goal is any expression that tells where the Theme ends up, or would end up, as a result of the motion. <i>We arrived in Paris before midnight.</i> Although always conceptually present and specific, Goal may sometimes be understood from context, rather than expressed by any separate constituent. <i>Our visitors arrived yesterday.</i> <i>私は画廊を出てようやくノーザンパランド・ホテルにたどりついた。</i>

Core Unexpressed

©Keio University



Being developed by Interns from Ecole Centrale de Lyon

Current Grants

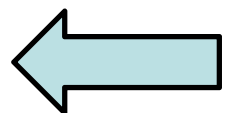
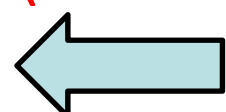
- Constructicon Building
 - Grant-in-Aid for Scientific Research
 - 2012-2015
- Full Text Annotation
 - Grant from the National Institute of Japanese Language and Linguistics
 - 2012-
- JFN Data Release
 - Grant-in-Aid for Scientific Research
 - 2013-2017



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)
- GoiTaiki (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) 



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_believe 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. <Experienter> *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 omou 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* *kangaeru*

NOM

ACC

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

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*] *kangaeru*

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.

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.'



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

- Giving frame

- *ageru.v*

- { [**Donor** .EXT.NP.*ga*]
[Recipient .Dep.NP.*ni*]
[Theme .Obj.NP.*o*]}]

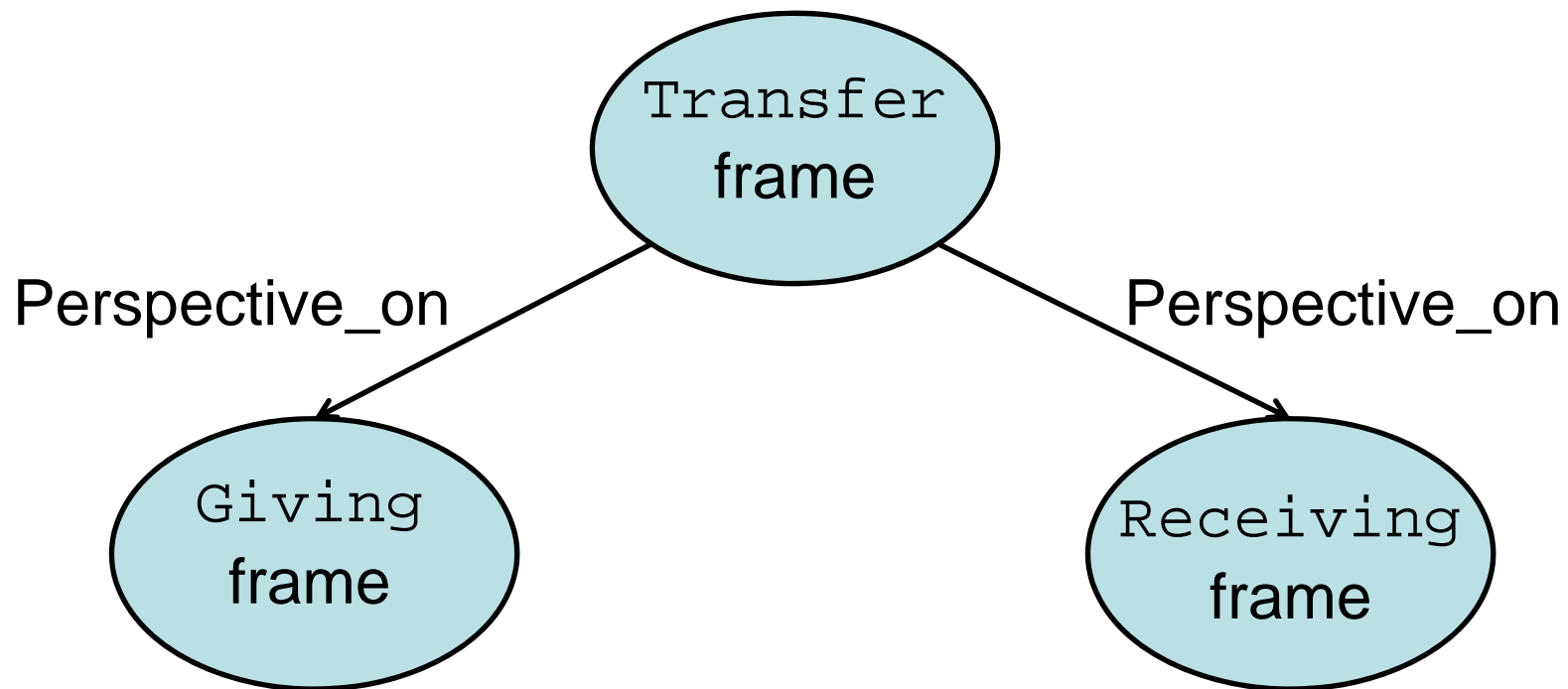
- Receiving frame

- *morau.v*

- { [Recipient .EXT.NP.*ga*]
[**Donor** .Dep.NP.*ni*]
[Theme .Obj.NP.*o*]}]



Frame-to-Frame relations between Giving & Receiving frames



⇒ 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)



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>

