

Constructicon Building in Japanese FrameNet: What it reveals about English and Japanese frames and constructions



JAPANESE FRAMENET

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



24th June, 2013
ICLC
University of Alberta, Canada

Outline

1. Introduction
2. Interlingual usefulness of frames
 - Coverage
 - Adequacy
3. Why we need a construction
4. Conclusions



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

Japanese FrameNet

Research questions:

- To what extent are the existing English-based semantic frames applicable to characterizing Japanese words?
- To what extent are the Frame-semantic and Construction Grammar approaches suitable for analyzing the Japanese lexicon and grammar?



Interlingual Usefulness of Frames



Interlingual Usefulness of Frames: Coverage

Applicability of **Existing FN frames for English** to full text annotation of Japanese texts

- “Core” data in the Book genre of BCCWJ corpus
- Full Text Annotation (running texts)
- Annotate all content words except for named entities
 - Verbs, Adjectives, Adverbs, and Nouns in 810 sentences

Tokens 87% (4,000 tokens/4,587 tokens)

cf. Types 85%

cf. Tokens in FN 69%



Japanese words 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 takeru – 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

Interlingual Usefulness of Frames: Coverage

- FN has been mainly annotating Verbs, Adjectives, and event nouns, but NOT conjunctions and nouns
- FN still needs to create frames for general words such as *experience*, *publication*, *account for*, and *take up*.
- We don't always need to create frames for culture-specific words

e.g. tatami – straw mat, syoozi – sliding paper,
husuma – sliding door, kyookaku – knight of the
town



=> Existing FN frames are useful in annotating
Japanese texts

Interlingual Usefulness of Frames: Adequacy

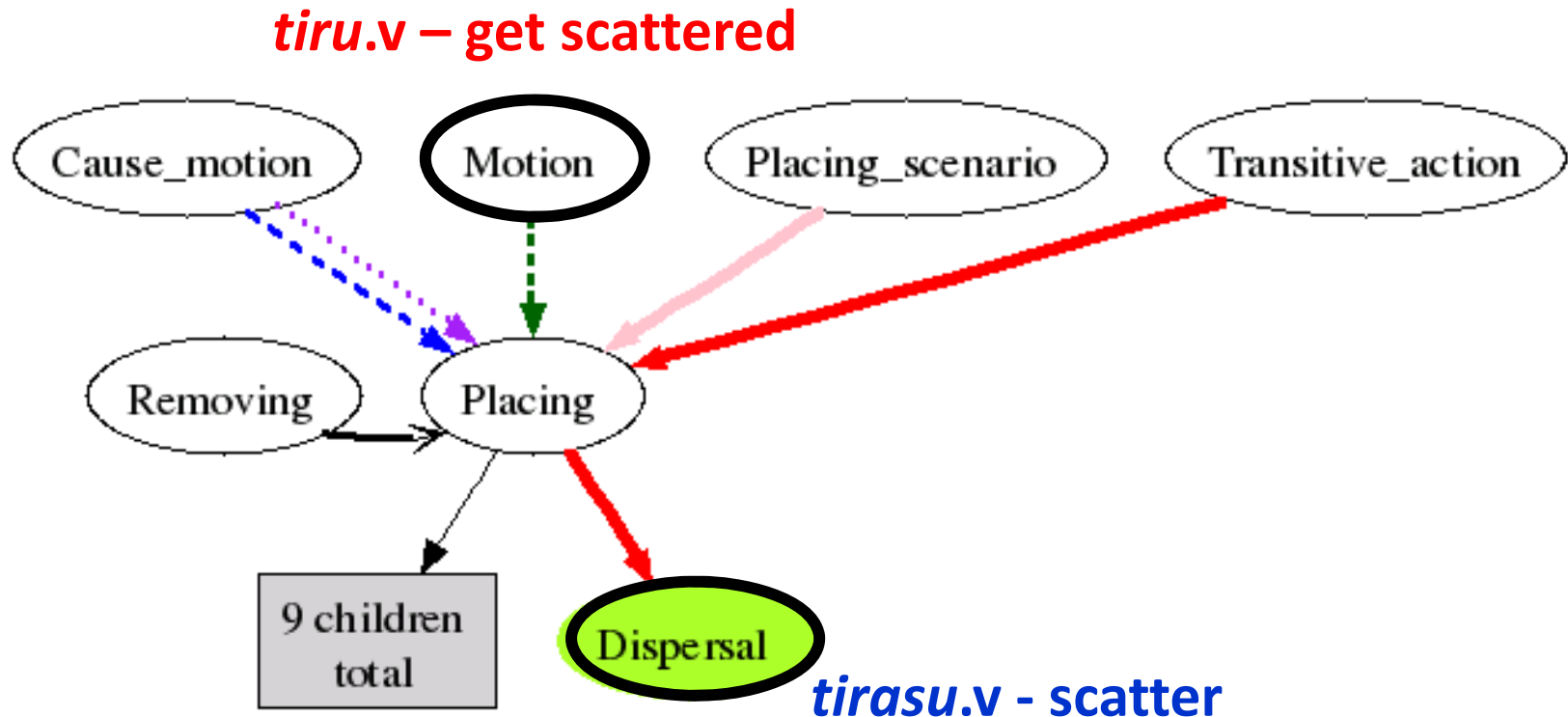
- Many pairs of intransitive-transitive pairs of verbs in J
 - Morphologically related

=> Some frames are missing in the current FN

- a. *sakura* *no hanabira ga* *tiru* Motion
 cherry.blossom GEN petals NOM get.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.’

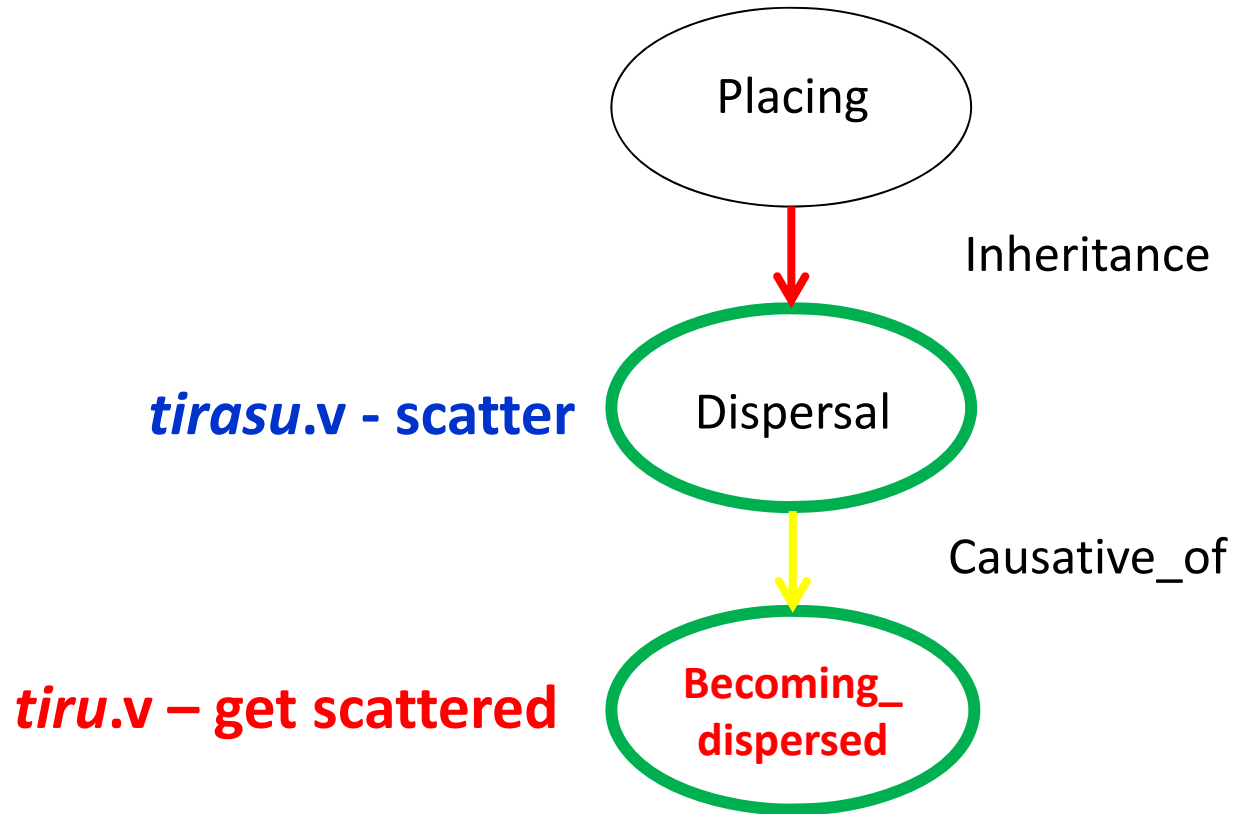


Frame-to-Frame Relations pertaining to Motion and Dispersal frames



=> Does not capture the relation between the pair of verbs

Proposal: Create Becoming_dispersed frame



=> For J, we need the
Becoming_dispersed frame, but do we also
need this frame for E?



Interlingual Usefulness of Frames: Adequacy

- Few frames in the current FN database with **intransitive** perspective
 - Become_silent, Become_triggered, Becoming, Becoming_a_member, Becoming_attached, Becoming_aware, Becoming_detached, Becoming_dry, Becoming_separated, Becoming_visible, and Expansion frames
- Few pairs of frames in the current FN database defined from **intransitive/inchoative** and **transitive** perspectives
 - Exception: Becoming_detached frame
 - Detaching frame
 - Being_detached frame
 - Filling frame
 - Fullness frame



Interlingual Usefulness of Frames

- Coverage
 - The current FN frame database has good coverage for full text annotation in Japanese
- Adequacy
 - Few frames in the current FN database with **intransitive** perspective
 - Few pairs of frames in the current FN database defined from **intransitive/inchoative** and **transitive** perspectives
 - ⇒ Does this mean that FN needs to create these frames?
 - ⇒ Does this mean that intransitive perspective is not often lexicalized in English?



Why we need a constructicon



Things which cannot be handled in JFN

- Anything other than “relations of predication, modification, and complementation”
- Linguistic objects “that function *as units* while at the same time having a describable internal structure”

(cf. Fillmore et al. 2012:312-313)

a. Complex particle

kiku tokoro ni yoru to 787 wa mondai rasii

hear place LOC depend QUOTE TOP problem seem

literal. ‘Depending/Based on (the) place (where I) hear, 787 seems problematic.’

= ‘Judging from (what I) hear, 787 seems problematic.’



Things which cannot be handled in JFN

b. 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.’

c. Clause-level construction

kore (no hoo) ga are yori nagai

This GEN side NOM that than long

‘This is longer than that.’

Constructional Analysis: Case 1

Hearsay construction

- The *Hearsay* construction indicates that the speaker has an *Attitude* about the *Proposition* attributed to a *Medium*.
- Construct Elements: *Medium, Proposition, Attitude*

a'. [_{Medium} *kiku*] [_{CEE} *tokoro ni yoru to*]
[_{Proposition} *787 wa mondai*][_{Attitude} *rasii*]

‘Judging from (what I) hear, 787 seems problematic.’

Construction evoking a frame: Case 1

Hearsay construction evokes the
Attribute_information frame:

- A PROPOSITION is attributed to a SPEAKER or a TEXT.

a'' *kiku* [_{FEE} *tokoro ni yoru to*]

[_{PROPOSITION} *787 wa mondai*] *rasii*

SPEAKER : DNI

'Judging from (what I) hear, 787 seems
problematic.'



Constructional Analyses: Case 2

Resultant_state construction

- The *Resultant_state* construction describes a *State* after an *Event* pertaining to an *Entity* has occurred.
- Construct Elements: *Entity, Event, State*

b' [_{Entity} *koi kiri ga*] *numa no ue ni*

[_{State= CEE} [_{Event} *orite*] [*iru*]

'A thick fog has fallen over the mire.'



Construction evoking a frame: Case 2

Resultant_state construction evokes the *State_continue* frame:

- The ENTITY remains in the specified STATE.

b'' [ENTITY *koi kiri ga*] *numa no ue ni*

[STATE=FEE *orite* [*iru*]]

'A thick fog has fallen over the mire.'



Constructional Analyses: Case 3

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*

c' [_{Entity1} *kore (no hoo) ga*] [_{Entity2} *are*] [_{CEE} *yori*]
[_{Feature} *nagai*]

'This is longer than that.'

Construction evoking a frame: Case 3

Comparative_inequality construction evokes the *Comparison_inequality* frame:

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



Why we need a constructicon

- Linguistic objects that function *as units* with a describable internal structure cannot be dealt with in JFN, but can be described in JFN constructicon
- The framework of construction annotation seems to be working for Japanese



Summary

- Interlingual usefulness of Frames

- Coverage

- The existing FN frame database has good coverage to be used in full text annotation of Japanese

- Adequacy

- Relative scarcity of frames with **intransitive** perspective in FN

⇒ Lack of coverage?

⇒ Intransitive perspective NOT often lexicalized in English?

- Why we need a constructicon



Linguistic objects that function *as units* with an internal structure can be described in JFN constructicon

Conclusions

- To what extent the existing English-based frames are applicable to characterizing Japanese words
 - Yet to find out whether the organization of frames of Japanese is different from that of English
- To what extent the Frame Semantic and Construction Grammar approaches are suitable for analyzing the Japanese lexicon and grammar
 - So far the Frame Semantic and Construction Grammar approaches seem suitable for analyzing Japanese



Selected References

- Baker, Collin. (2006). "Frame Semantics in Operation: The FrameNet Lexicon as an Implementation of Frame Semantics" ICCG4.
- Fillmore, Charles J. (2006). "The Articulation of Lexicon and Construction" 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 Lingüística Aplicada.
- Fillmore, Charles J. (2013). "Berkeley Construction Grammar." In Hoffman, Thomas & Graeme Trousdale (Eds.), *The Oxford Handbook of Construction Grammar*, 111- 132. Oxford University Press.
- Fillmore, Charles J. and Collin F. Baker (2009) "A Frames Approach to Semantic Analysis". In Heine, Bernd and Heiko Narrog (Eds.), *The Oxford Handbook of Linguistic Analysis*, 313-339. Oxford University Press.
- Fillmore, Charles J., Russell R. Lee-Goldman, and Russell Rhodes. (2012). "The FrameNet Construction". In Boas, Hans C. and Sag, Ivan A. (Eds.), *Sign-based Construction Grammar*, 309-372. Center for the Study of Language and Information.
- Hasegawa, Yoko, Russell Lee-Goldman, Kyoko Ohara, Seiko Fujii, and Charles J. Fillmore. (2010). "On expressing measurement and comparison in Japanese and English". In Boas, Hans C. (Ed.) *Contrastive Construction Grammars*, 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. (2009). "Frame-based contrastive lexical semantics in Japanese FrameNet: The case of *risk* and *kakeru*". In Boas, Hans C. (Ed.), *Multilingual FrameNets in Computational Lexicography: Methods and Applications*, 163-182. Mouton de Gruyter.
- Ohara, Kyoko Hirose (2012). "Semantic Annotations in Japanese FrameNet: Comparing Frames in Japanese and English". LREC2012.
- Ohara, Kyoko Hirose (2012). "Japanese FrameNet: Toward Construction Building for Japanese". Plenary Talk, ICCG7.
- Ruppenhofer, Joseph, et al. (2010). "FrameNet II: Extended Theory and Practice."

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>

