

Browsing Japanese FrameNet with FrameSQL

Hiroaki SATO

Senshu University Office #9513
2-1-1 Higashimita, Tamaku, Kawasakishi
Kanagawaken, JAPAN
thc0408@gmail.com

Abstract

The Japanese FrameNet project at Keio University is building a Japanese lexicon that records the valence descriptions of Japanese words. A part of its data is now accessible at the FrameSQL website which the author created ten years ago and has been updating ever since. FrameSQL not only shows a clear view of the headword's grammar and combinatorial properties of the Japanese lexicon, but also relates the Japanese lexicon with its counterparts in English, Spanish and German within the framework of Frame Semantics. FrameSQL puts together the four separate lexicons, and users can access them seamlessly, as if they were a unified database. Mutual hyperlinks among these databases and the new search mode make it easy to compare the semantic structures of corresponding lexical units across the languages, and they could be useful for building multilingual lexical resources, or more generally for multilingual studies.

1 Introduction

The Japanese FrameNet project (JFN) of Keio University (<http://jfn.st.hc.keio.ac.jp/index.html>) aims at building a lexicon that records the valence descriptions of Japanese words (Ohara et al., 2003), based on Frame Semantics (Fillmore, 1976) and corpus data. JFN works in collaboration with the Berkeley FrameNet project (BFN), which is building an English lexicon (Baker et al., 1998; Fillmore et al., 2002) at the International Computer Science Institute in Berkeley, California (<http://framenet.icsi.berkeley.edu/>).

berkeley.edu/).

FrameSQL (Sato, 2003) is a web-based application that the author originally designed ten years ago to search the BFN database, a lexicon of contemporary written English that the project is producing for both computational and lexicographic interests. The author has incorporated into FrameSQL the Japanese lexicon of JFN, the Spanish lexicon (Subirats and Sato, 2004) of the Spanish FrameNet project (SFN, <http://gemini.uab.es:9080/SFNsite/>), and the German lexicon (Erk et al., 2003) of the Saarbrücken Lexical Semantics Acquisition Project (SALSA, <http://www.coli.uni-saarland.de/projects/salsa/page.php?id=index>). FrameSQL offers the same user-interface for searching these lexical data.

FrameSQL can search and view the JFN data released in March of 2009 with a standard web browser. You do not need to install any additional software tools to use FrameSQL, nor do you even need to download the JFN data to your local computer, because FrameSQL accesses the database of the server computer, and executes searches. The server computer handles the JFN data with MySQL, a popular relational database application that can execute complex searches. Directly searching a MySQL database requires learning the database query language SQL, but with FrameSQL, you do not have to learn SQL at all. You only select several search parameters and a form of the output search results on the web browser, and FrameSQL translates these parameters into their corresponding SQL command and executes the search.

2 Basic Idea

FrameNet is based on the theory of Frame Semantics. The central idea of the theory is that word meanings must be described in semantic **frames** which schematically represent concep-

tual structures. Each semantic frame has a group of **lexical units** (LUs). A lexical unit is a pairing of a word with a sense (Cruse, 1986) whose semantic properties are described with semantic roles called **frame elements** (FEs). For example, the *Arriving* frame has a list of LUs such as *approach.N*, *approach.V*, *arrive.N* and *arrive.V*, and semantic properties of example sentences in this frame are described with its own set of FEs such as *Theme* (the object that moves) and *Goal* (the place where *Theme* ends up).

BFN makes an inventory of frames and FEs for annotating English sentences. Most of them work for annotating sentences of other languages, because frames are meant to characterize conceptual structures at a basic level of description. In the case of the *Arriving* frame, the corresponding Japanese LUs are *たどりつく.V* (reach) and *着く.V* (arrive), and JFN basically uses the same FEs to annotate Japanese example sentences taken from the KOTONOHA corpus developed by the National Institute for Japanese Language (http://www.kokken.go.jp/en/research_projects/kotonoha/kotonoha/). FrameSQL uses lists of frames, LUs and FEs as search parameters.

3 Frame Search Mode

There are several search modes in FrameSQL. Figure 1 shows the Frame search mode for the JFN data (<http://sato.fm.senshu-u.ac.jp/jfn21/notes/index.html>), where you can search and view annotated sentences of a single frame at a time.

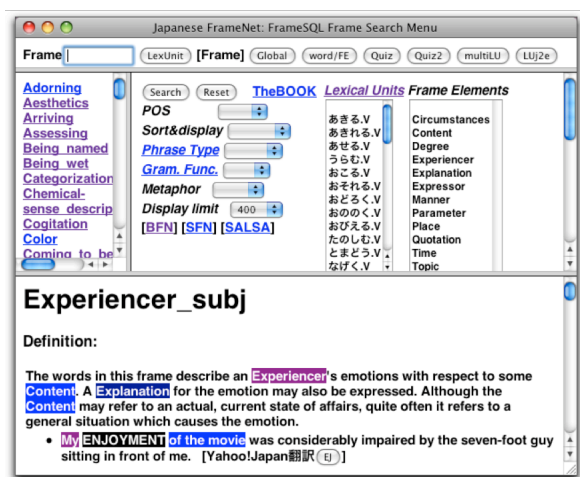


Figure 1. Frame Search Mode

The startup menu of the Frame search mode is the search menu of the *Experiencer_subj* frame, which is evoked by an LU describing an *Experiencer's*

emotions with respect to some *Content*. *Experiencer* and *Content* are FEs used in this frame, and definitions of the frame and its FEs are displayed in the lower pane of Figure 1. You can select search menus of other frames by clicking a frame name in the list of the middle-left pane.

Figure 2 shows the Frame search mode of the *Experiencer_subj* frame where a user selects *怖がる.V* (be scared) in the LU list. When the user clicks on the Search button, FrameSQL executes the search and shows a summary of search results of this LU in the lower pane, as in Figure 2. The search summary table shows a set of FEs used in each annotated sentence. You can access the annotated sentence by clicking on the hyperlinked number on the left of the table. For example, when the user clicks on the hyperlinked number on the lower pane, the corresponding annotated sentence appears on the middle-right pane, as in Figure 3.

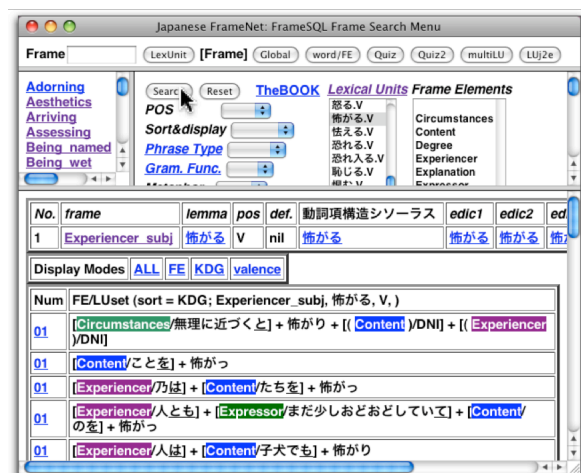


Figure 2. Search of LU 怖がる.V

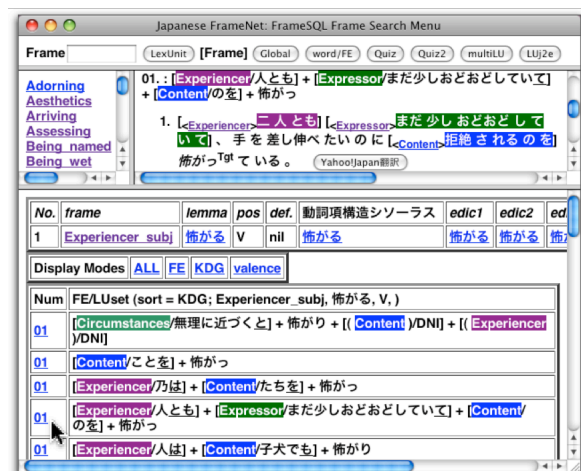


Figure 3. Showing annotated sentences

The annotated sentence on the middle-right pane of Figure 3 contains FEs *Experiencer*, *Expressor* and *Content*, as in (1).

- (1) [*<Experiencer>*二人とも]
 [*<Expressor>*まだ少しおどおどして
 いて]、手を差し伸べたいの
 に [*<Content>*拒絶されるのを]
 怖がっ^{Tgt}ている。

There are hundreds of FEs and frames defined in the English lexicon of BFN. Users often need to see those definitions when they browse annotated sentences. FrameSQL puts hyperlinks on FEs of annotated sentences so that users can access FE definitions just by clicking on the links. For example, when the user clicks on the hyperlink of *Expressor*, the definition of the FE *Expressor* appears on the lower pane, as in Figure 4.

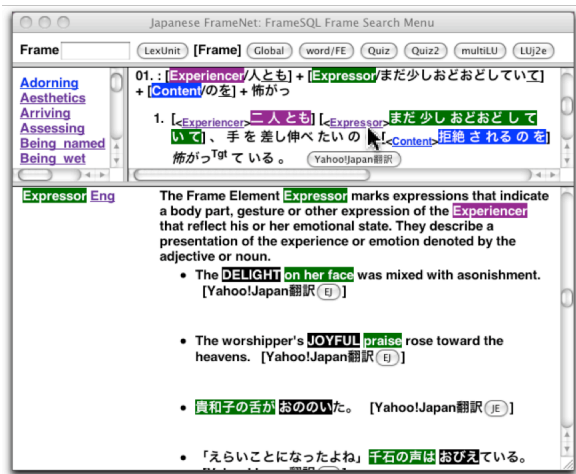


Figure 4. Accessing FE definitions

4 Mutual Hyperlinks of the Same Frame

At the left-bottom of the middle-right pane in Figure 1, there are three hyperlinks: [BFN] [SFN] [SALSA]. They are linked to FrameSQL search menus of BFN, SFN and SALSA of the same *Experiencer_subj* frame respectively. FrameSQL search menus for these projects are basically the same. Once you get accustomed to one of those search menus, you can search those lexical databases in the same manner, and compare LUs of the same frame among those languages.

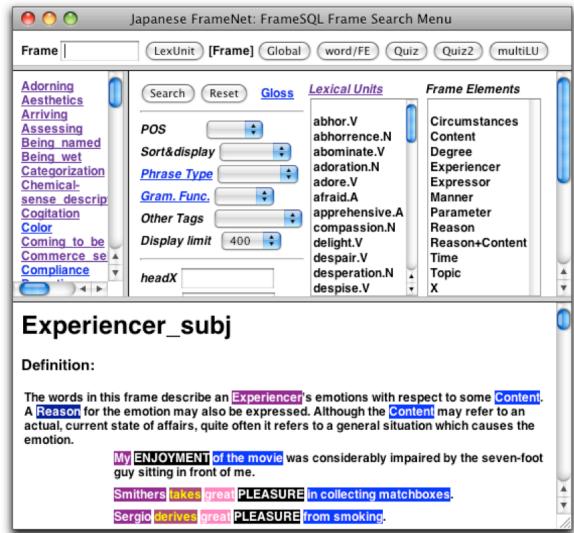


Figure 5. BFN search menu



Figure 6. SFN search menu

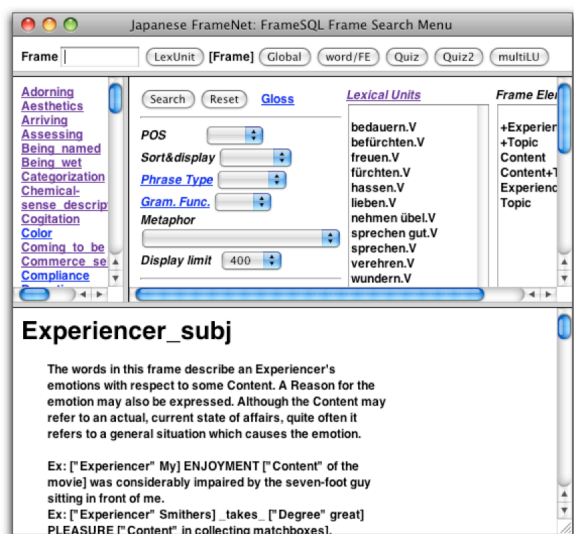


Figure 7. SALSA search menu

For example, when you click on [BFN] of Figure 1, the corresponding BFN search menu of the *Experiencer_subj* frame appears in the middle-right pane, as in Figure 5. The user-interface for English lexicon in Figure 5 is basically the same as that for Japanese lexicon in Figure 1. The user can access the SFN search menu for Spanish lexicon (Figure 6) and the SALSA search menu for German lexicon (Figure 7) of the same *Experiencer_subj* frame in the same manner.

5 Bilingual LU menu

Another search mode of FrameSQL makes it possible to compare semantic structures of corresponding LUs of JFN and BFN. In this mode, FrameSQL shows a list of LUs of JFN and BFN belonging to the same frame. For example, Figure 8 shows the bilingual LU list of the *Motion_directional* frame, where the FE *Theme* moves in a certain *Direction* which is often determined by gravity or other natural, physical forces.



Figure 8. Bilingual LU menu

When the user selects 落ちる.V and *drop.V* and presses the Search button, the summary of search results appears as in Figure 9. When the user clicks on a hyperlinked number in the table, the corresponding annotated sentences appear on the web browser.

FEset	落ちる.V	drop.V
	14	33
Goal + Path + Source + Theme		01
Goal + Path + Theme		02
Goal + Source + Theme	01	
Goal + Theme	07	11
Path + Theme	01	07
Source + Theme	01	08
Theme	03	03
(Goal) + Source + Theme	01	
(Path) + Theme		01

Figure 9. Bilingual summary table

6 Concluding Remarks

FrameNet uses hundreds of frames and FEs to annotate corpus data, and it sometimes is difficult for users to understand them, when they browse the database. FrameSQL is useful because users can easily access their definitions by clicking hyperlinked frame and FE names. FrameSQL puts together four separate lexical databases, and users can access them seamlessly, as if they were a unified database. Mutual hyperlinks among these databases and the new search mode make it easy to compare the semantic structures of corresponding LUs across languages, and it could be useful for building multilingual lexical resources, or more generally for multilingual studies.

References

- Baker, C.F., Fillmore, C.J., and Lowe, J.B. (1998). The Berkeley FrameNet project. In *Proceedings of the COLING-ACL*, pp.86--90, Montreal, Canada.
- Cruse, D.A. (1986). *Lexical Semantics*. Cambridge: Cambridge University Press.
- Erk, K., Kowalski, A., Pinkal, M. (2003). A Corpus Resource for Lexical Semantics. In *Proceedings of the Fifth International Workshop on Computational Semantics (IWCS)*, pp.106--121, Tilburg, The Netherlands.
- Fillmore, C.J. (1976). Frame Semantics and the Nature of Language. In *Annals of the New York Academy of Sciences: Conference on the Origin and Development of Language and Speech*, Volume 280, pp.20--32.
- Fillmore, C.J., Baker, C.F., Sato, H. (2002). The FrameNet Database and Software Tools. In *Proceedings of the Third International Conference on Language Resources and Evaluation*, pp.1157--1160, European Language Resources Association.
- Ohara, K.H., Fujii, S., Saito, H., Ishizaki, S., Ohori, T., Suzuki, R. (2003). The Japanese FrameNet Project: A Preliminary Report. In *Proceedings of Pacific Association for Computational Linguistics (PACLING'03)*, pp. 249--254. Halifax, Canada.
- Sato, H. (2003). FrameSQL: A Software Tool for FrameNet. In *ASIALEX '03 Tokyo Proceedings*, pp.251--258, Asian Association of Lexicography, Tokyo, Japan.
- Subirats, C., Sato, H. (2004). Spanish FrameNet and FrameSQL. In *Proceedings of Building Lexical Resources from Semantically Annotated Corpora*, pp.13--16, European Language Resources Association.