How FrameSQL Shows the Japanese FrameNet Data

Hiroaki Sato

Senshu University Office #9513

2-1-1 Higashimita, Tamaku, Kawasaki, Kanagawa, JAPAN

E-mail: thc0408@gmail.com

Abstract

FrameSQL is a web-based application which the author (Sato, 2003; Sato 2008) created originally for searching the Berkeley FrameNet lexical database. FrameSQL now can handle the Japanese lexical database built by the Japanese FrameNet project (JFN) of Keio University in Japan. FrameSQL can search and view the JFN data released in March of 2009 on a standard web browser. Users do not need to install any additional software tools to use FrameSQL, nor do they even need to download the JFN data to their local computer, because FrameSQL accesses the database of the server computer, and executes searches. FrameSQL not only shows a clear view of the headword's grammar and combinatorial properties of the database, but also relates a Japanese word with its counterparts in English. FrameSQL puts together the Japanese and English lexical databases, and the user can access them seamlessly, as if they were a unified database. Mutual hyperlinks among these databases and the bilingual search mode make it easy to compare semantic structures of corresponding lexical units between these languages, and it could be useful for building multilingual lexical resources.

1. Introduction

The Japanese FrameNet project (**JFN**) of Keio University (http://jfn.st.hc.keio.ac.jp/index.html) in Japan has been building a lexicon that records valence descriptions of Japanese words (Ohara et al., 2003). 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 (http://framenet.icsi.berkeley.edu/). FrameSQL is a web-based application which I (Sato, 2003; Sato 2008) created originally for searching the Berkeley FrameNet lexical database.

FrameSQL has been updated and now it can handle the JFN lexical database. FrameSQL can search and view the JFN data released in March of 2009 on a standard web browser. Users do not need to install any additional software tools to use FrameSQL, nor do they even need to download the JFN data to their 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, users do not have to learn SQL at all. They 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.

FrameSQL not only shows a clear view of the headword's grammar and combinatorial properties of the JFN lexical data, but also relates a Japanese word with its counterparts in English. This paper compares the Japanese 悩ます.*V* (to annoy) with its English counterparts *trouble.V, annoy.V, puzzle.V* and *perplex.V.*, and FrameSQL shows how it displays their semantic structures on a web browser.

2. Basic Idea

FrameNet is based on the theory of Frame Semantics (Fillmore, 1976). The central idea of the theory is that word meanings must be described in semantic **frames** which schematically represent conceptual 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 *Experiencer_obj* frame describes a situation in which some phenomenon provokes a particular emotion in an experiencer. This frame has a list of LUs such as *annoy.V*, *scare.V* and *terrify.V*. Semantic properties of example sentences in this frame are described with its own set of FEs such as *Experiencer* (the one that reacts emotionally or psychologically) and *Stimulus* (the event or entity which brings about the emotional or psychological state of the *Experiencer*).

3. Mutual Links between BFN and JFN

FrameSQL puts several hyperlinks between the BFN and JFN data. These links are useful for comparing these two lexical databases. They are also useful for those who have a little knowledge of the Japanese language. Those users will find FrameSQL less difficult to use, if they start with searching the BFN English data with FrameSQL. The search menus of FrameSQL for both the data are basically the same, and once users get accustomed to the English search menus, they can use the Japanese search menus in the same way. So I will start with explaining how to use FrameSQL for searching the BFN lexical data.

There are several search modes in FrameSQL. Figure 1 shows the **LexUnit** search mode (http://sato.fm.senshu-u.ac.jp/fn2_14alpha/notes/index2.html) for searching the BFN data, where users can search and view annotated sentences of a single LU at a time.

Users can select another search mode by clicking a button on the upper pane. When they click on the [Frame] button on the upper pane, they can use the **Frame** search mode, where they can search and view annotated sentences of a single frame at a time.
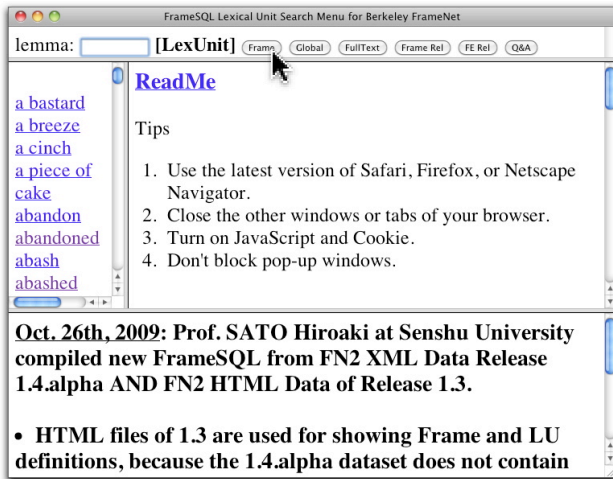
Figure 1: LexUnit Search Mode for BFN

3.1. Frame Search Mode for BFN

FrameSQL uses lists of frames, LUs and FEs as search parameters. Figure 2 shows the Frame search mode where a user selects

- *Experiencer_obj* in the frame list of the middle-left pane,
- *annoy.V* in the LU list of the middle-right pane,
- *Stimulus* in the FE list of the middle-right pane, and
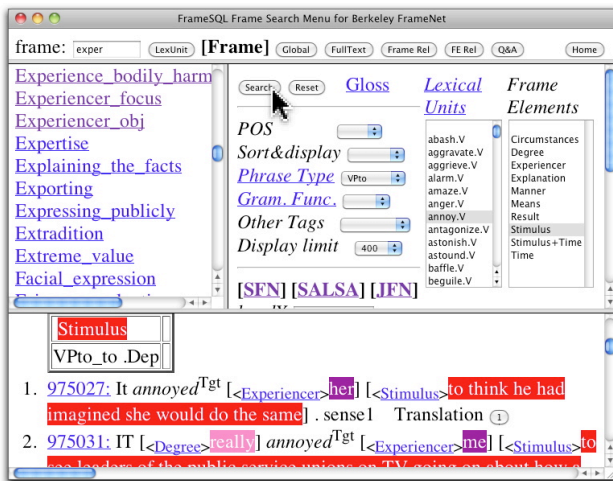- *VPto* in the Phrase Type list of the middle-right pane.

Figure 2: Frame Search Mode for BFN

When the user clicks on the [Search] button in the middle-right pane, FrameSQL translates these search parameters into a SQL command, and executes the search.

In the lower pane of Figure 2, FrameSQL shows search results which include all the example sentences with the FE *Stimulus* assigned to the phrase type to-infinitival verb phrase of the LU *annoy.V* in the *Experiencer_obj* frame of the BFN data.

At the bottom of the middle-right pane in Figure 2, there are three hyperlinks: [SFN] [SALSA] [JFN]. They are linked to FrameSQL search menus of Spanish

FrameNet (Subirats and Sato 2004), the Saarbrücken Lexical Semantics Acquisition Project (Erk et al. 2003) and JFN respectively. FrameSQL search menus for these projects are basically the same. Once users get accustomed to one of those search menus, they can search those lexical databases in the same manner, and compare LUs of the same frame among those languages.

3.2. Frame Search Mode for JFN

When a user clicks on the [JFN] hyperlink in Figure 2, the search menu for the same *Experiencer_obj* frame of the JFN data appears on the middle-right pane, as in Figure 3.
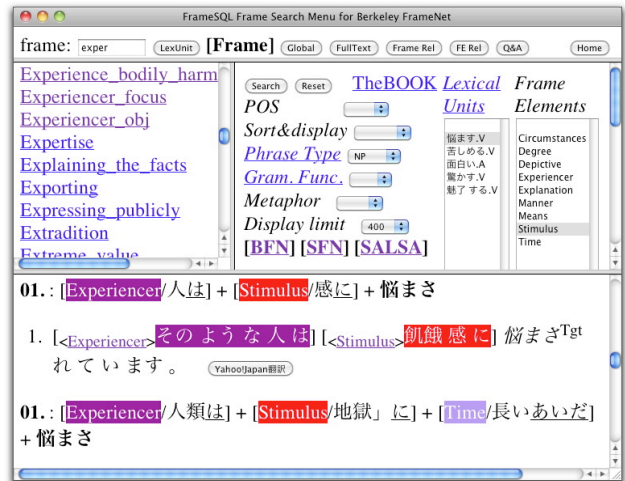

Figure 3: Frame Search Mode for JFN

In Figure 3 the user selected 悩ます. *V* (to annoy) in the LU list and NP in the Phrase Type list of the middle-right pane, and executed the search. The lower pane of Figure 3 shows search results which include all the example sentences with the FE *Stimulus* assigned to the phrase type NP of the LU 悩ます. *V* in the *Experiencer_obj* frame of the JFN data.

FrameSQL adds translation buttons to the end of each annotated sentence for translating the Japanese sentence into English. When a user clicks on the translation button in Figure 3, FrameSQL sends the Japanese sentence to the Yahoo translation website (http://honyaku.yahoo.co.jp/transtext), and the translation website shows the translation of the annotated sentence, as in Figure 4.
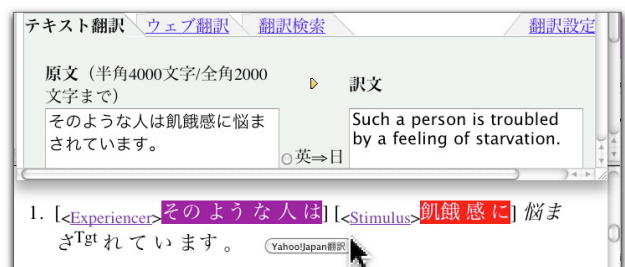

Figure 4: Yahoo translation

FrameNet uses fine-grained semantic frames and FEs to characterize meanings and uses of words. Since most of FEs are defined in terms of specific semantic frames, there are hundreds of FEs used in annotations. As a result, it sometimes is difficult for users to understand all the FEs, when they browse the databases. FrameSQL is helpful because the user can easily access FE definitions by clicking hyperlinked frames and FE names.

Figure 5 shows how a user looked up the definition of the FE *Stimulus* in the lower pane by clicking on the hyperlinked FE name in the annotated sentence.
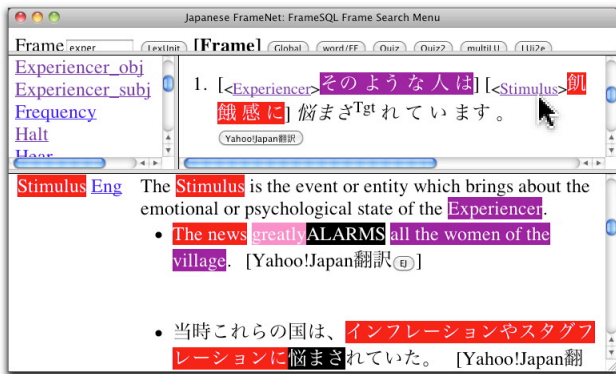

Figure 5: Accessing the FE definition

Figure 6 shows how a user displayed the definition of the *Experiencer_obj* in the lower pane by clicking on the hyperlinked frame name in the middle-left frame list.
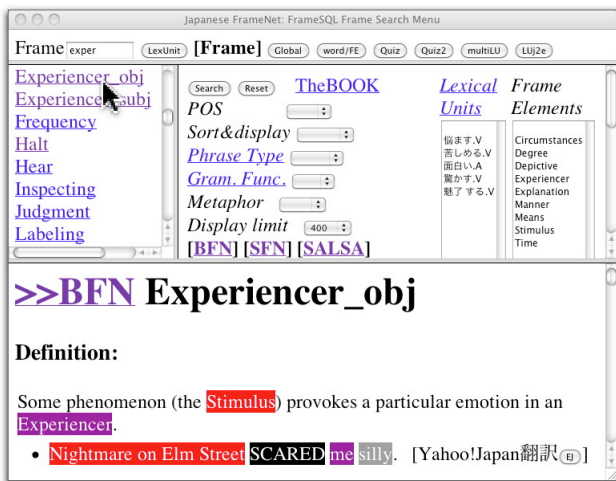

Figure 6: Accessing the frame definition

4. Bilingual LU menu

Another search mode of FrameSQL enables users to compare semantic structures of corresponding LUs of JFN and BFN data. In this mode, FrameSQL shows a table of FEs assigned to annotated sentences. For example, Figure 7 shows how FEs of the *Experiencer_obj* frame are assigned to annotated sentences of the Japanese LU 悩ます.V and its English translation equivalents *trouble.V*, *annoy.V*, *puzzle.V* and *perplex.V*.
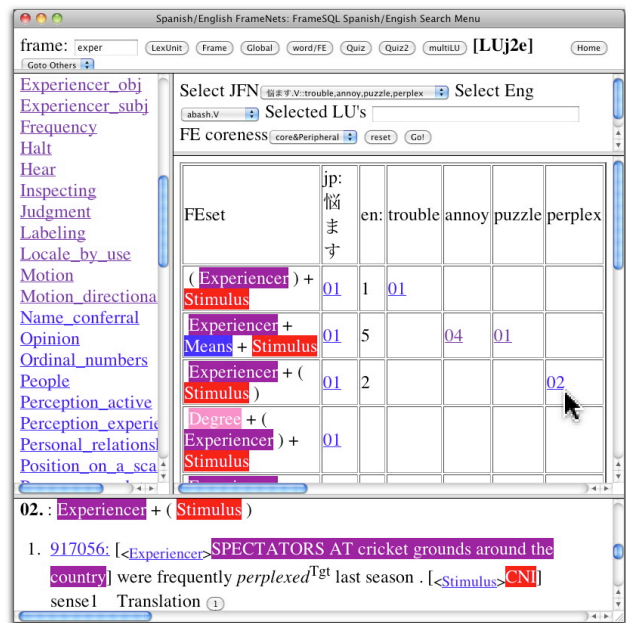

Figure 7: Bilingual Search Mode

The numbers in the table of the middle-right pane show occurrences of example sentences annotated with colored FEs. For example, the fourth row *Experiencer + Stimulus 01 ... 02* shows that the LU 悩ます.V has one annotated sentence with the FEs *Experiencer* and *Stimulus*, and *perplex.V* has two English sentences with the same FEs. The numbers in the table are linked with the corresponding annotated sentences. The user can access Japanese and English annotated sentences by clicking on the hyperlinked numbers. In Figure 7, the user clicked on the hyperlinked number 02 of *perplex.V* in the fourth row, and FrameSQL displayed sentences which contain this FE set in the lower pane.

FrameSQL compares translation equivalents in a Japanese-to-English electronic dictionary and LUs of BFN for making a list of corresponding LUs. For example, translation equivalents of 悩ます.V in the electronic dictionary are as follows.

(1) trouble, worry, bother, annoy, puzzle, persecute, fret, harass, perplex

The LU 悩ます.V belongs to the *Experiencer_obj* frame of JFN, and BFN registered the following English LUs in the same frame.

(2) abash.V, aggravate.V, aggrieve.V, alarm.V, amaze.V, anger.V, annoy.V, antagonize.V, astonish.V, astound.V, baffle.V, beguile.V, bewilder.V, bewitch.V, ... startle.V, stimulate.V, sting.V, stir.V, stun.V, stupefy.V, surprise.V, terrify.V, thrill.V, tickle.V, torment.V, traumatize.V, trouble.V, unnerve.V, unsettle.V, upset.V, vex.V, vexation.N, wow.V

FrameSQL extracts the English LUs which appear both in (1) and (2), and pairs 悩ます. *V* with the English LUs *trouble.V*, *annoy.V*, *puzzle.V* and *perplex.V*. FrameSQL makes English translation equivalents for every Japanese LU in the same manner, as shown in Table 1. This table is used for comparing semantic structures of corresponding LUs in Figure 7.

Frames	Japanese LUs	English LUs
Experiencer_obj	悩ます	trouble, annoy, puzzle, perplex
Experiencer_obj	苦しめる	torment, trouble, distress
Experiencer_obj	面白い	∅
Experiencer_obj	驚かす	surprise, astonish, amaze, startle
Experiencer_obj	魅了する	fascinate, charm, enchant

Table 1: Corresponding JP an EN LUs

5. Conclusion

FrameSQL puts together JFN and BFN lexical databases, and the user can access them seamlessly, as if they were a unified database. Mutual hyperlinks among these databases and the bilingual search mode make it easy to compare semantic structures of corresponding LUs between these languages, and it could be useful for building multilingual lexical resources, or more generally for multilingual studies.

References

Baker, C.F., Fillmore, C.J., 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, Las Palmas, Spain.

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.

Sato, H. (2008). New Functions of FrameSQL for Multilingual FrameNets. In *Proceedings of the Sixth Language Resources and Evaluation Conference*, European Language Resources Association, Marrakech, Morocco.

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, Lisbon, Portugal.