FrameNet as an Assessment Tool for English-to-Japanese Translation

Yoko Hasegawa
University of California, Berkeley

In collaboration with the late Professor Charles Fillmore and Russell Lee-Goldman, I recently published in the journal *Constructions and Frames* an article "On the Universality of Frames: Evidence from English-to-Japanese Translation." We demonstrated that the frames created in the Berkeley FrameNet project for the annotation of English texts can also serve as an assessment tool for the accuracy of English-to-Japanese translation. This idea of FrameNet as a translation testing tool is not unique to ours, however. In recent years, some interesting articles proposing a similar idea applying to other languages have been published.

Today, I am going to demonstrate how we analyzed translations by applying FrameNet frames and how such an analysis helps us compare similar constructions across languages.

Before discussing FrameNet as a translation-assessing tool, I would like to reminisce a little. As most of you certainly know, Professor Fillmore was fluent in Japanese. He acquired this competence while he was stationed in the U.S. Army in Kyoto right after WWII. His duty in the Army was to intercept coded Russian conversations on short-wave radio. He told us that in those days, Russian coded communication consisted solely with numbers, so he learned only those Russian words corresponding to the numbers from 0 to 9. He loved to demonstrate in Russian how such communication sounded like.

When he was off duty, he walked around the streets in Kyoto with a notebook in hand and conducted fieldwork on the Japanese language. He especially liked to go to coffee shops and to listen to customers' casual conversations. After he was discharged, he taught English at a local girls' junior high school in Kyoto. He was extremely popular among those young girls, and he said he fell in love with 80% of them. While teaching English, he took classes at Kyoto University, where he learned canonical classical Japanese literature. He would proudly recite particularly the opening of *Taketori Monogatari* (*The Tale of the Bamboo Cutter*).

He returned to the United States to receive his doctorate at the University of Michigan, and then he spent ten years teaching at Ohio State University before he came to UC Berkeley in 1971. Professor Fillmore started offering his Japanese Linguistics Seminar in 1987, when I was a graduate student, and continued offering it until the summer of 2012, when his health had deteriorated. I was extremely fortunate to have such a great scholar-teacher for such a long period of time.

Translation is generally understood as an art rather than a science. Passages that are anchored deeply in the source-text author's culture are bound to lose some meanings due to different expectations and background experience on the part of the target-text readers. In our article, we

were concerned exclusively with scientific writings because scientific writing is a genre in which factual accuracy and conceptual clarity are mandatory, while considerations of aesthetic elegance and cultural nuances are normally less relevant. Therefore, we expected scientific translations to be a base-line testing ground for the utility of FrameNet. If the frames needed for scientific texts turn out not to be cross-linguistically applicable, there will be little reason to expect FrameNet frames to serve as a common platform of comparison for translations in such areas as social structure, religion, or art.

Frames are schematic representations of types of events, situations, individuals, and things, including the participants, props, parts, and their relations to each other and to the larger situations. Words evoke frames, and words are understood with a particular frame as background. The FrameNet database groups words with the same background knowledge into frames, and it defines these frames and the parts of the frame, called *frame elements*. Some frames can be language-specific; however, most of them are expected to be universal. We argued in the article that if the semantic structure of an original text is carefully analyzed in terms of the frames evoked by its constituent words and the ways in which the elements of those frames are realized, such frames and frame elements, as well as their interconnections, should somehow be retained by accurate translations.

This is a reasonable expectation because in FrameNet, many kinds of morphosyntactic idiosyncrasies are abstracted away from the basic frame structures. That is, the same semantic information can be expressed by using different parts of speech. For example, *to decide* and *to make a decision* evoke the same frame. We have shown in the article that FrameNet frames are quite versatile even when applied cross-linguistically to languages that prefer different event-encoding strategies.

In Translation Studies, several methods to test semantic accuracy of translations have been proposed. However, they invariably entangle highly subjective judgements. We believed that FrameNet can be a tool that enables more objective assessment, by providing the frames according to which translation addition and omission are identified. Furthermore, FrameNet analyses enable us to compare similar constructions across languages.

In the *Constructions and Frames* article, we concentrated on sentences involving causation because the differences in rhetorical preference between English and Japanese are said to manifest most significantly in causal expressions. When Situation A causes Situation B, both situations are likely to be expressed in terms of clauses in Japanese, but Situation A is frequently represented by an abstract NP in English. For example:

(1) SitA: Diagnostic methods of Parkinson's disease have advanced.
     SitB: Experts became aware that Parkinson's disease can attack those younger than 40.

(2) Jp: (Backtranslation) Because [diagnostic methods advanced]$_{SitA}$, [experts have become aware that Parkinson's disease can attack those younger than 40]$_{SitB}$.
     Eng: [Advancement of the diagnosis methods]$_{SitA}$ has made [experts aware that Parkinson's disease can attack those younger than 40]$_{SitB}$.

In the Japanese version, both events are expressed by clauses, whereas the causing event in English is expressed as the NP "advancement of the diagnosis." The direct translation that maintains the English constructional equivalent is also possible in Japanese. However, it would sound unnatural to Japanese ears.

I now illustrate how we applied FrameNet frames to the actual analysis of English-to-Japanese translation. One of the English sentences we considered is:

(3)   Better diagnosis has made experts aware that Parkinson's disease can attack those younger than 40.

In this sentence, the matrix predicate is *made*, which evokes the `Causation` frame:

> `Causation` frame: A CAUSE causes an EFFECT. Alternatively, an ACTOR, a participant of a (implicit) CAUSE, may stand in for the CAUSE.

This frame yields this annotation:

(4)   Better diagnosis$_{\text{CAUSE}}$ has MADE experts aware that Parkinson's disease can attack those younger than 40$_{\text{EFFECT}}$

The corresponding Japanese translation maintains the `Causation` frame by the use of the postpositional phrase *ni.yotte* 'due to':

(5)   *Shindan hōhō ga shinpo-shita*$_{\text{CAUSE}}$ *koto*   *NI.YOTTE*
        diagnostic methods advanced   due to

        40-sai.miman demo pākinson-byō o hasshō-suru koto ga wakatte.kita$_{\text{EFFECT}}$
        even those who are under 40 can have symptoms of Parkinson's disease

        (Backtranslation) DUE TO the fact that diagnostic methods advanced$_{\text{CAUSE}}$, we've become aware that even those who are under 40 can have symptoms of Parkinson's disease$_{\text{EFFECT}}$

We can see frame matching here. Next, *better* evokes the `Progress` frame:

> `Progress` frame: An ENTITY changes from a PRIOR_STATE to a POST_STATE in a sequence leading to improvement.

Diagnosis is identified as the ENTITY of this `Progress` frame:

(6)   BETTER diagnosis$_{\text{ENTITY}}$

In English, the `Progress` frame can be evoked by *better*, but denoting an event by an adjective is not possible in Japanese. Therefore, this event needs to be evoked by such a lexical unit as *shinpo-shita* 'progressed':

(7) *Shindan hōhō ga*ENTITY *SHINPO-SHITA* *koto ni.yotte* …
     diagnostic methods   advanced

     (Backtranslation) Due to the fact that diagnostic methodsENTITY ADVANCED …

Aware evokes the `Awareness` frame:

`Awareness` frame: A COGNIZER has a piece of CONTENT in their model of the world.

This results in this frame-element assignment:

(8) expertsCOGNIZER AWARE that Parkinson's disease can attack those younger than 40CONTENT

In Japanese, the `Awareness` frame can be evoked by *wakaru* 'understand'; the COGNIZER frame element is implicit here:

(9) *40-sai.miman demo pākinson-byō o hasshō-suru*CONTENT
     even those who are under 40 can have symptoms of Parkinson's disease

     koto ga *WAKATTE*.kita
                aware

     (Backtranslation) We've become AWARE that even those who are under 40 can have symptoms of Parkinson's diseaseCONTENT

Finally, we note the `Attack` frame:

`Attack` frame: An ASSAILANT physically attacks a VICTIM (which is usually but not always sentient), causing or intending to cause the VICTIM physical damage.

This frame is used metaphorically to express how Parkinson's disease affects people:

(10) Parkinson's diseaseASSAILANT can ATTACK those younger than 40VICTIM

Here, we need to consider a major rhetorical difference between English and Japanese, namely the topic-worthiness hierarchy phenomenon. Other things being equal, this hierarchy of topic-worthiness is normally observed in Japanese:

Human (first/second person) > Human (third person) > Animate Nonhuman > Inanimate

Therefore, sentences like *Parkinson's disease can attack people*, in which a non-human entity is selected as the subject (occupying a more salient syntactic position) and a human is downgraded as the object are likely to be restructured in such a way that the human occupies the subject position in the Japanese translation. This requires a frame perspective shift from `Attack` to `Getting_disease`:

`Getting_disease` frame: A VICTIM starts off without the DISEASE, and then comes to suffer from it.

---

(11) *40-sai.miman demo*_{VICTIM}   *pākinson-byō o*_{DISEASE}   *HASSHŌ-SURU*
Those younger than 40   Parkinson's disease   acquire

(Backtranslation) Those younger than 40_{VICTIM} can ACQUIRE Parkinson's disease_{DISEASE}

The `Attack` and `Getting_disease` frames are related, but not closely. Connecting these two frames systematically is a challenging and yet very interesting area to explorer in future FrameNet investigation. These frames can be associated by the following frame-to-frame relations. The `Attack` frame (in its literal sense) is connected via the inheritance relation to the `Transitive_action` frame, and in turn to the `Objective_influence` frame. The `Objective_influence` frame is connected to the more general frame of `Influence` through a perspective-on relation. The other perspective on the `Influence` frame is the `Be_influenced` frame. This is the frame from which the `Getting` frame inherits, and the `Getting_disease` frame is a subtype of the `Getting` frame.
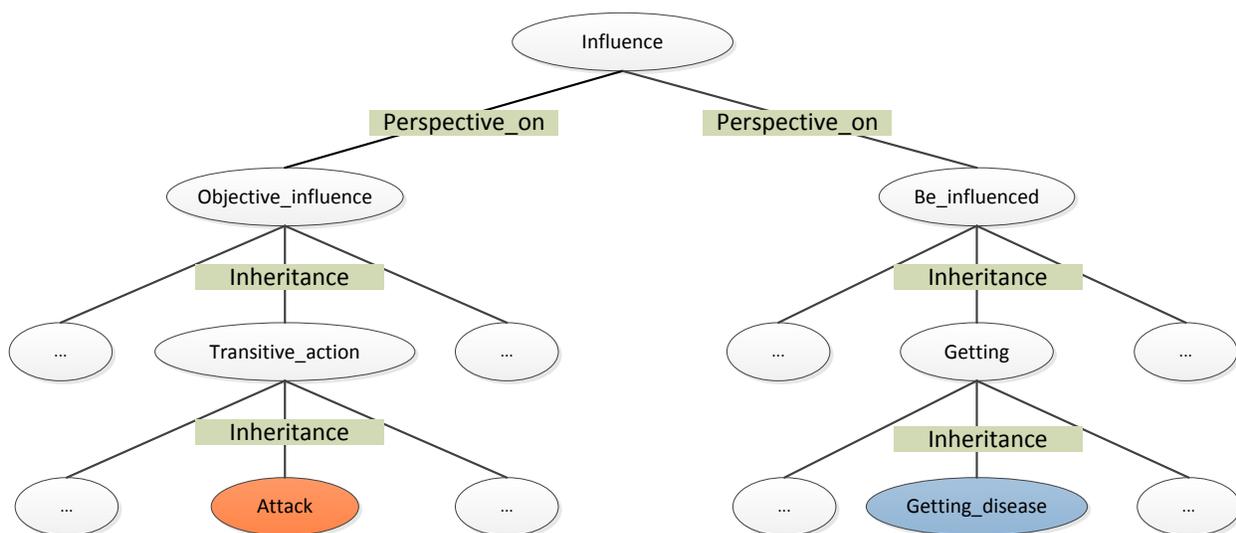
As we have seen, these two sentences superficially differ considerably,

(12) Better diagnosis has made experts aware that Parkinson's disease can attack those younger than 40.

(13) (Japanese translation backtranslated) Due to the fact that diagnostic methods advanced, we've become aware that even those who are under 40 can have symptoms of Parkinson's disease.

However, all major frames and frame elements of the source English text are encoded in its Japanese translation; therefore, according to the FrameNet's frame test, this translation is judged highly accurate.

As demonstrated, a full-fledged frame-semantic account, with FrameNet as a core component, provides a detailed specification of the meaning of a sentence. Consequently, such an analysis provides a useful dimension within which to understand how to compare similar constructions cross-linguistically. Such an analysis also highlights similarities between seemingly dissimilar constructions and intrigues avenues in cross-linguistic constructional analysis. This potential power of FrameNet in semantic investigation has been recognized widely, and several languages have developed their own FrameNet database. I now turn to the Japanese FrameNet and demonstrate its utility in Japanese-to-English translation assessment.

The sentence I would like to considered is:

(14) *Saibā-kūkan ga   saibā-kyōi ni   taishi   zeijakuna   baai ni wa, jōhō no*
      cyberspace   cyber threats   against vulnerable occasion   information

      jiyūna   ryūtsū no   kakuho ga   konnan ni   naru.
      free   flow   maintenance difficult   become

      (Backtranslation) When cyberspace is vulnerable to cyber threats, the maintenance of free flow of information becomes difficult.

The matrix clause of this sentence is:

(15) *jōhō no jiyūna ryūtsū no kakuho ga konnan ni naru.*

      (Backtranslation) the maintenance of free flow of information becomes difficult.

The matrix predicate of this sentence is *naru* 'become', which evokes the `Becoming` frame:

> `Becoming` frame: An ENTITY ends up in a FINAL_STATE or FINAL_CATEGORY which it was not in before.

(16) *jōhō no jiyūna ryūtsū no kakuho ga*ENTITY   *konnan ni*FINAL_STATE   *NARU*
      the maintenance of free flow of information   difficult   become

      (Backtranslation) the maintenance of free flow of information ENTITY BECOMES difficult FINAL_STATE

The `Becoming` frame has a CIRCUMSTANCE as its non-core frame element, which marks the set of conditions under which the ENTITY enters the FINAL_STATE or FINAL_CATEGORY.

(17) *Saibā-kūkan ga Saibā-kyōi ni taishi zeijakuna baai ni wa*_{CIRCUMSTANCE},
When cyberspace is vulnerable to cyber threats

*jōhō no jiyūna ryūtsū no kakuho ga*_{ENTITY}   *konnan ni*_{FINAL_STATE}   *NARU*_{ENTITY}
the maintenance of free flow of information   difficult   become

(Backtranslation) When cyberspace is vulnerable to cyber threats_{CIRCUMSTANCE}, the maintenance of free flow of information_{ENTITY} BECOMES difficult_{FINAL_STATE}

This sentence is translated by a professional translator as:

(18)   Vulnerability of cyberspace to cyber threats may hamper free flow of information.

The matrix predicate of this translation is *hamper*, which evokes the `Hindering` frame:

`Hindering` frame: A HINDRANCE makes it more difficult for a PROTAGONIST to complete their intended ACTION.

A FrameNet account of this sentence is:

(19)   Vulnerability of cyberspace to cyber threats_{HINDRANCE} may HAMPER free flow of information_{ACTION}

This is an interesting instance of Japanese-to-English translation. We have observed that when Situation A causes Situation B, both situations are normally expressed as clauses in Japanese, but Situation A is frequently represented as an abstract NP in English. Here, even the event that is encoded as a CIRCUMSTANCE in Japanese is promoted to the subject of the English translation. This frame correspondence is schematically represented as follows (LU: lexical unit):

Jp: [CIRCUMSTANCE]   [ENTITY]   [FINAL_STATE]   `Becoming` LU
                                                            negative

Eng: [HINDRANCE]   `Hindering` LU   [ACTION]

This generalization can predict that a sentence like

(20) *Dojō no hiyokudo ga teika suru to*_{CIRCUMSTANCE},
If soil nutrients diminish

*shokubutsu ga kikō hendō o kanwa suru chikara ga*_{ENTITY}   *yowaku*_{FINAL_STATE}   *NARU*
plants ability to ease climate change   weak   become

(Backtranslation) If soil nutrients diminish_{CIRCUMSTANCE}, plants ability to ease climate change_{ENTITY} BECOMES weak_{FINAL_STATE}

will be translated as:

(21)   Limited soil nutrients_{HINDRANCE} may HAMPER plants ability to slow climate change_{ACTION}

When the FINAL_STATE is perceived positively, the English translation will utilize a lexical unit that evokes an antonymous frame, e.g. `Assistance`:

`Assistance` frame: A HELPER benefits a BENERITED_PARTY by enabling the culmination of a GOAL that the BENERITED_PARTY has.

Jp: [CIRCUMSTANCE]   [ENTITY]   [FINAL_STATE]   `Becoming` LU
                                  positive

Eng: [HELPER]   `Assistance` LU   [GOAL]

To conclude, a full-fledged frame semantic analysis provides a detailed specification of the meaning of a sentence, and I have shown the crucial role of FrameNet, by identifying the major frames encoded in the source text and scrutinizing whether they reappear in the translation.

FrameNet reveals certain types of differences between idiomatic English and Japanese, as in preferences regarding the expression of event causation. When straightforward frame matching results in conflict with a stylistic norm of Japanese, a paraphrase of the source text has been translated. For example, where English spoke of a disease attacking the victim, Japanese spoke of the victim acquiring the disease.

By comparing the frames evoked by major constituents of a text and its translation, we are able to assess translation accuracy more objectively than would have been possible with hitherto proposed translation evaluation methods.