

WHEN NONCE WORDS BEHAVE LIKE REAL WORDS

A CASE STUDY OF THE JAPANESE VERB OSO(WARER)U

KOW KURODA *

KEIKO NAKAMOTO **

HITOSHI ISAHARA *

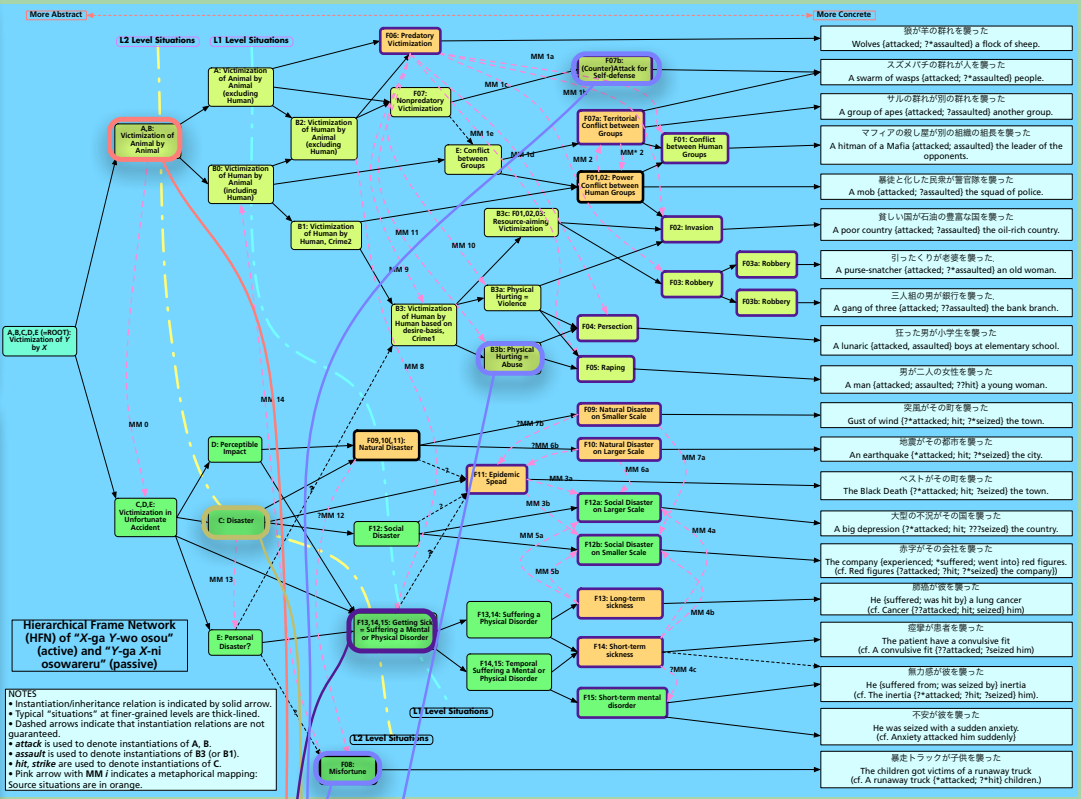
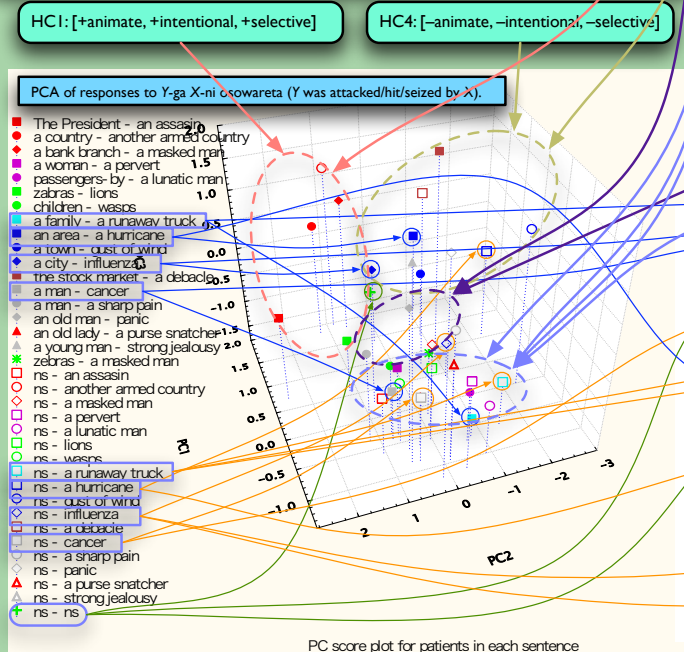
*NATIONAL INSTITUTE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY (NICT), JAPAN

**BUNKYO UNIVERSITY, JAPAN

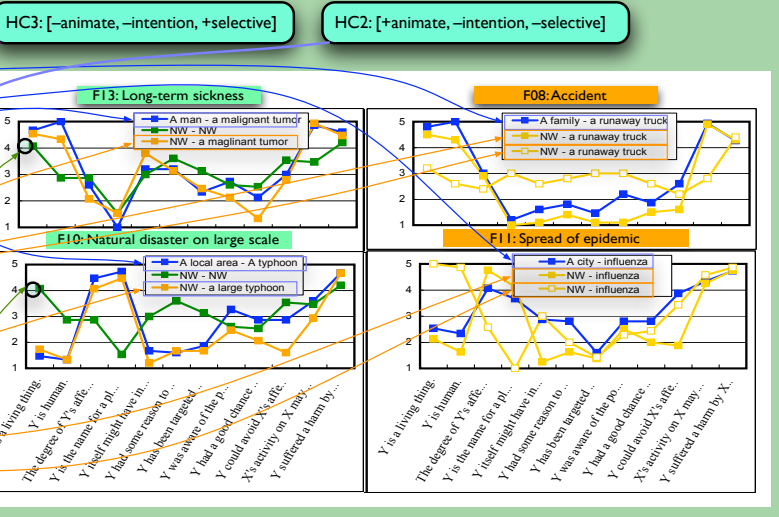
Goals: We want to
1. show the effectiveness in characterizing semantic interpretation (not only as co-composition of lexical meanings but also) as having the property of **attraction to the most conceivable situation**, adopting the Frame Semantics view that conceivable situation is something that the (co-composed) meaning of a sentence is "evaluated" against.
2. show the usefulness of the "vector space" approach to the meaning representation at sentential and/or phrasal level as well as at lexical level.

Method
1. Manually construct a Hierarchical Semantic Frame Analysis (HFNA) of the situations of **harm-causation class**, based on the analysis of a Japanese verb.
2. Determine the set of semantic features/characteristics that account for the HFNA. In this study, we concentrated on investigation into attraction-to-situation effects by representing the expected meaning of a nonce word in a situation-evoking context.
3. Conduct a psychological experiment (Semantic Feature Rating; SFR) to evaluate the analysis/hypothesis.

Results
 Our predication was basically confirmed in that, in the context **victim(Y)** was attacked/hit/seized by **harm-causer(X)** (**victim(Y)-ga harm-causer(X)-ni osowareta** in Japanese),
1. The meaning of a nonce word for **X** and **Y** are predictable from the HFNA if the context evokes a particular situation strongly enough (selection among candidate is assumed).
2. Different nouns have different strengths of evocation, depending on the degrees of their representativeness for a potential situation-specific semantic role/frame element.
3. In general, the names of (typical instances of) harm-causer have stronger attraction effects than those of victim. (typicality of victimhood is rather hard to define).
4. By and large, we found 4 classes of harm-causer: **HC1, HC2, HC3, HC4**.
5. While the locations of F03b, F07b and F08 are not well-predicted, others situations have nice, predicted locations in the vector space.



Class	ID	English translation of rated feature	Rated feature in Japanese
Harm-causer	1	X is a living thing.	Xは生き物である。
Harm-causer	2	X chose Y for its target.	XはYを選んで襲った。
Harm-causer	3	X is visible.	Xは目に見える実体である。
Harm-causer	4	X couldn't help doing it to Y.	XがYを襲ったのは仕方ないことだった。
Harm-causer	5	X is human.	Xは人間である。
Harm-causer	6	X had an aim to do so.	Xは目的を持って襲った。
Harm-causer	7	X is a natural phenomenon.	Xは自然現象である。
Harm-causer	8	X did so to satisfy its desire or needs.	Xは自分の欲求を満たすために襲った。
Harm-causer	9	X planned to take off something from Y.	XはYから何かを奪うつもりだった。
Harm-causer	10	X is the name for a sickness.	Xは病名である。
Harm-causer	11	X's activity can kill Y.	XはYを襲って死なせることがある。
Harm-causer	12	X is a collection of living things.	Xは生き物の集まりである。
Victim	1	Y is a living thing.	Yは生き物である。
Victim	2	Y had a good chance to prepare for X's activity.	YがXの襲撃に備えるのは簡単だった。
Victim	3	Y had some reason to be victimized by X.	YにはXに襲われる何らかの理由があった。
Victim	4	Y is human.	Yは人間である。
Victim	5	Y was aware of being victimized by X.	YはXに襲われる可能性に気づいていた。
Victim	6	X's activity on X may cause X to die.	YはXに襲われたのが原因で死ぬことがある。
Victim	7	Y is the name for a place.	Yは場所を表す名前である。
Victim	8	Y could avoid X's harm.	YはXに襲われるのを回避することもできた。
Victim	9	The degree of Y's affectedness is greater than the individual scale.	Yの襲われ方は個人/個体の規模を超える。
Victim	10	Y suffered a harm by X's activity on Y.	YはXに襲われて実際に被害を受けた。
Victim	11	Y has been targeted by X long before.	YはXに以前から狙われていた。
Victim	12	Y itself might have invited X's activity on it.	YがXに襲われたのにはYにも責任がある。



PCA score plot for patients in each sentence