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
3-5 Hikari dai Seika cho, Soraku gun, Kyoto, 612-0289
**Bunkyo University, Japan 3337 Minami Ogishima, Koshigaya, Saitama, 343-811

1 Introduction

It can never be false to say that the meaning of word w in a specific context $C(w)$ results from the complex interaction of w's “lexical” meaning $m(w)$ and the meaning of $\bar{m}(w) = m(C(w))$, i.e., the “meaning of its context.” No lexicographic work can be done without assuming the distinction between m and $\bar{m}$, but it is not at all obvious how to make such a decision because we hardly know what $\bar{m}(w)$ really is. This is, in part, what makes it difficult to tell exactly what $m(w)$ really is. We approached to this problem experimentally, hoping to forge a connection from the theory of language to real human behaviors.

Based on psychological experiments, called the task of **semantic feature rating** (SFR) on a Japanese verb, we do two things in this paper. First, we show that the meanings assigned to nonce words in specific contexts are predictable if we suppose that semantic interpretation is situationally based, as claimed by Frame Semantics (FS) [4] and Berkeley FrameNet (BFN) [6, 19] and if we are able to specify, say in the form of a lattice, the hierarchical system of situations for which candidate sentences are interpreted. We suggest that the situational view could lead to a successful specification of how **co-composition** [17] is constrained. Second, we pose the question of how contextually induced complex meanings are constructed. Note that such meanings can be specified not only for constituents like “___ NP-*wo* V,” namely VP of Japanese, but also for nonconstituents like “NP-*ga* ___ V”.[1] How this is handled theoretically is an open question, but we are skeptical of whether an account based on “movements at LF” is a valid account of it since there is no guarantee that an account of this type fits the observed human behavior unless LF movements are proved to be “real.” [NOTE explain LF]

1.1 The basic idea

If human understanding is, as FS claims and BFN assumes, **situationally based**, it follows that:

[1] Traditionally, the structure of VP of Japanese is characterized as $[_{\text{IP}}$ NP[+nom] $[_{\text{VP}}$ NP[−nom] V]]. In our example, $[_{\text{IP}}$ NP-*ga* $[_{\text{VP}}$ NP-*wo* V]] is an instance of this template under the condition that -*ga* and -*wo* serve as nominative and accusative markers, respectively.

(1) **Interpretation as selection:** the interpretation of a sentence $s = w_1 w_2 \cdots w_n$ ($W(s) = \{ w_1, w_2, \ldots, w_n \}$) is not simply “constructed” from the set of **lexical** meanings $\{ m_1, m_2, \ldots, m_n \}$ (m_i is the meaning of w_i) but is given as a “selection” from a predetermined set of possible situations for which s can be interpreted.

(2) **Attraction-to-situation (A-to-S) effect:** Given this selectional property, interpretations of a given sentence s are expected to be “attracted” to a particular situation, and word sense modulation arises as a “side effect” of this attraction.

This predicts the following:

(3) A-to-S is effective even if all arguments of a predicate (“governor” in the sense of FrameNet [6]) are not explicitly given as long as “frame-evoking elements,” which do not need to be words and can be collocational units, evoke frames strongly enough.

We tested this prediction (3) experimentally and obtained positive results. To this aim, we used the SFR technique [15]. In an SFR task, roughly, participants are asked to rate a word or phrase within a sentence for a pre-determined set of (usually fine-grained) semantic features or characteristics that seem to be necessary to fully account for the interpretational variation of S. Details of the SFR technique are presented in §2.3.

1.2 Why the latent semantics of nonce words?

As pointed out above, a semantic description of any lexical item, say a word w, presupposes an appropriate discrimination of w's lexical meaning from the meaning it gains from its context $C(w)$, i.e., so-called “contextual effects.” There is no guarantee, however, that good discrimination can be achieved on all occasions because, in principle, there is no way to do it. We would like to say that this is the “dark side” of co-composition [17]. At present, all we can do is rely almost entirely on the intuition of lexicographers and linguists. We hope that our research into the latent semantics of nonce words[2)] will contribute to the investigation of a systematic account of contextual effects.

1.3 A theory of semantic attraction

1.3.1 Comparison of "constructivist" and "selectional" theories

Most theories of semantic interpretation, e.g., Generative Lexicon Theory (GLT) [17], are "constructivist" ones in that the meaning of a complex unit (e.g., phrase and sentence) is constructed from the lexical meanings of its "parts." This is the traditional view of the meaning of construction. However, another kind of model is conceivable. We may refer to a "selectional" theory (as in the Darwinian theory of evolution). Let us begin by examining what will happen if semantic interpretation is "selectional" rather than purely constructive in nature.

One of the best examples of a selectional theory of semantic interpretation would be *Optimality Theory* (OT) [2, 16]. It is selectional in that it characterizes the interpretation of a given sentence s as the selection of an "optimal" interpretation. An optimal interpretation is the interpretation that wins out of a set of "candidate" interpretations generated in some way.

Note that, under this selectional view, semantic interpretation need not be truly compositional. The component for candidate generation, usually called GEN in the OT literature, may need to be compositional, whereas the component for output evaluation, usually called EVAL, cannot be. What the evaluation component does is select the one best candidate. In OT, this is implemented by a "ranking mechanism." Candidates generated by GEN are "scored" against a set of "constraints" and ranked according to their scores.

A more radical model is conceivable, however. Note that even GEN need not be compositional when candidate generation is done by enumeration. We interpret Frame Semantics (FS) as implementing such a radical model in a sense to be explained later.

It is well-known that GLT argues against the so-called "sense enumerative lexicon." But it is not clear what happens if we conceive of a database that enumerates all the situations for which all sentences are interpreted. We investigate this in some detail below.

1.3.2 Test of the selectional view of interpretation

We interpret FS as another, more radical selectional theory, since FS allows words and phrases in a discourse to freely "evoke" frames independently of each other. In a radical interpretation, there is no requirement for structure building to occur.

Based on this, we can hypothesize the following:

(4) Possible semantic interpretations of a given sentence S are "attracted" to (ideally) one of the most likely situations.

If this prediction is correct, then a nonce word w^* should be feature-rated very much like a real word if the context of its occurrence $C(w^*) = W(s) - w^*$ (meaning word sequence except w^*) "evokes" a specific situation strongly enough. We tested this hypothesis through psychological experiments using sentences containing *osou*. As explained in §2.1, the Japanese verb *osou* is a rather polysemous verb. Its English translations include *attack*, *hit*, and *seize* (see Appendix 2.1 for relevant details). The setting for our experiments is explained below using English analogs.

In our experiments, we used cases such as those in (5) for $C(w^*$ for $\langle$ victim $\rangle)$ and cases such as those in (6) for $C(w^*$ for $\langle$harm-causer$\rangle)$:

(5) ___ { a. was attacked by; b. was hit by; c. was seized by; d. suffered from } ⟨ harm-causer ⟩ (or ___ suffered ⟨harm⟩)

(6) ⟨ victim ⟩ { a. was attacked by; b. was hit by; c. was seized by; d. suffered (from) } ___.

Our prediction will be confirmed if SFRs for nonce words in C_1 and C_2 conditions are interpreted like real words in C_1, on the one hand, and if they are different from C_3, on the other:

(7) For the passive form "X-ga Y-ni osowareta,"

a. C_0: X is a real word for ⟨victim⟩; Y is a real word for ⟨harm-causer⟩ (Baseline 1)

b. C_1: X is a real word for ⟨ victim ⟩; Y is a nonce word for ⟨harm-causer⟩

c. C_2: X is a nonce word for ⟨ victim ⟩; Y is a real word for ⟨harm-causer⟩

d. C_3: X is a nonce word for ⟨ victim ⟩; Y is a nonce word for ⟨harm-causer⟩ (Baseline 2)

In (7b), the attraction effect of the word for ⟨ harm-causer ⟩ in the *oso(warer)u*-context can be detected. In (7c), the attraction effect generated by the word for ⟨ Victim ⟩ in the *oso(warer)u*-context can be detected. We tested this prediction using psychological experiments and obtained positive results.[3)] The results for C_1, C_2, and C_3 were obtained from different groups of participants.

1.4 Review of research into the "semantics of nonce words"

As far as we know, no intensive research into the "semantics of nonce words" has been attempted to date. One study [10] investigated the meaning of "syntactic frame/patterns" in the following way. Nonsensical sentences such as *The rom gorped the blickit to the dax*, *The grack mecked the zarg* were presented to participants, who were asked to rate the likelihood of various

2) We know this phrase sounds really like an oxymoron, but we do not know of any other term to express our concept. This might, we suspect, explain why this line of research is very rare.

3) To be precise, experiments on *osou*-contexts and *osowareru*-contexts were conducted on different occasions, so they are not directly comparable. This paper reports on the latter experiment.

semantic properties that could be true of the nonsense verbs in them. The results suggested that the syntactic frames encoded specific meanings, even if the verbs did not have lexical meanings. This experiment used the same technique as ours, but it had different goals and implications from our results.

2 Specifying "attractors" of interpretation

2.1 Semantics of *osou*

Let us briefly describe the relevant semantics of the Japanese transitive verb *osou* that we used in our experiment. It is a rather polysemous verb used to denote a wide range of situations or cases of *victimization* (but note that Japanese has distinct words for *victim*, i.e., *gisei-sha* (犠牲-者) and *higai-sha* (被害-者). Its English translations span over different classes of verbs. The overall picture can be seen from Figure 1. As easily seen, the meanings of *osou* and *osowareru* at the most abstract level are ⟨ harm-causer ⟩-ga ⟨ victim ⟩-wo osou (meaning "⟨ harm-causer ⟩ attacks/hits ⟨ victim ⟩") and ⟨ harm-causer ⟩-ni ⟨ victim ⟩-ga osowareru (both mean "⟨ victim ⟩ is/are attacked/hit by ⟨ harm-causer ⟩").

2.2 Identifying the situation lattice

The system of situations for which F1 and F2 are interpreted is represented by the lattice in Figure 2, which is called a hierarchical frame network (HFN). This was manually constructed from the corpus examples and validated through psychological experiments.

It should be noted, however, that it would not be appropriate to interpret the lattice in Figure 5 as a lattice of *osou*'s lexical meanings. A better interpretation would be that the HFN specifies the (partial) **ontology of harm or harm-causation** to which *osou*-sentences always refer. This interpretation was confirmed experimentally in [15].

2.2.1 How the HFN is related to "senses" of *osou*

Words senses are sensitive to granularity. This means that word sense definitions will make no sense unless they make reference to a level of granularity. The lowermost situations F01, F02, . . . , F15 would correspond to the finest-grained word senses. Most definitions for *osou* in Japanese lexica come between those two granularity levels. The top division between volitional subjects ({ A, B }) and nonvolitional subjects ({ C, D, E }) corresponds to the most basic division of sense differentiation. It is suggestive that *osou* can be translated as *attack* or *assault* for situations under { A, B }, whereas it cannot be translated in this way for situations under { C, D, E }.[4] For the latter, *hit* and *seize* are translations. In particular, *seize* is appropriate for situations under { F13, F14, F15 }, except for idiomatic cases like *panic attack* and *heart attack*.

The most coarse-grained distinction does not correspond to the distinction between the literal and metaphorical senses. Metaphorical senses appear all around the lattice, as indicated by links in magenta with an "MM*i*" index, where source and target domains are indicated.

It is reasonable to question if the situation/sense lattice for *osou*-sentences in Figure 5 has a wide enough coverage of *osou*-senses, if not exhaustive. Though indirect, we have two sources of evidence. First, the lattice is the result of a careful manual annotation/analysis of all instances (413 in total) of *osou*- or *osowareru*-sentences taken from a reasonably large corpus [20] of 500,000 Japanese-English pairs. For instances out of 413, 95% instances of the corpus data were successfully classified. We conducted another psychological experiment [15] to see to what degree the sense hierarchy is valid and obtained a positive result.

Another source is an informal study that found that, while the sense lattice in Figure 5 was constructed to account for the sense variation of *osou*-sentences, the lattice covered the sense variations of *gisei-sha*) and *higai-sha*), both meaning *victim* in English with different connotations.[5] Roughly, 80% of *gisei-sha* uses and 60% of *higai-sha* uses were covered, though precise evaluation has not been done yet. In this sense, we guess that the lattice in Figure 5 is not only a lattice of *osou*-sentences, but also a lattice of victimization situations in general.

2.2.2 HFN specifies units of selectional restrictions

It is reasonable to believe that the situations in the HFN correspond to the "units of selectional restrictions" on *oso(ware)u*-sentences in that each situation specifies a combination of finer-grained semantic roles and only a limited number of combinations are allowed for *oso(ware)u*-sentences. Possible combinations are as follows: (i) "⟨ natural disaster ⟩-ga ⟨ area ⟩-wo osou" (meaning "⟨ natural disaster ⟩ hit ⟨ area ⟩"), (ii) "⟨ man with mal-intention ⟩-ga ⟨ opponent ⟩-wo osou" (meaning "⟨ man with mal-intention ⟩ hit ⟨ opponent ⟩"), (iii) "⟨ robber ⟩-ga ⟨ bank ⟩-wo osou" (meaning "⟨ robber ⟩ attacked ⟨ storehouse of valuables ⟩"), and (iv) "⟨ social disaster ⟩-ga ⟨ domain of activity ⟩-wo osou" (meaning "⟨ social disaster ⟩ hit ⟨ domain of activity ⟩"). This is admittedly a strong claim, but it has been validated through psychological experiments reported in [14].

2.3 Background for SFR

The semantic feature rating (SFR) task, defined in [15], is an experimental procedure in (9), based on a (reasonable) theoretical assumption (8):

[4] One sense of *attack*, in the meaning of ⟨ accuse ⟩ and ⟨ criticize ⟩, based on metaphor, is systematically missing in the use of *osou*. Here, *kougeki(suru)* (攻撃 (する)), one of the hyponyms of *osou*, has a metaphorical sense. This verb refers to situations under A, namely F01, F02, F06, and F07, and it is hard to use this verb to refer to other situations.

[5] *gisei-sha* tends to refer to one or more victims who were seriously injured and dead, whereas *higai-sha* tends to refer to one or more victims who survived.

English verbs that translate OSOU	112 (TOTAL)	L0 = Sub L1 Level	L1	Semantic Classes at Level 1	L2	Semantic Classes at Level 2	L3	Semantic Classes at Level 3
attack[+human(s)]: rob	4	7	10	Resource-threatenig situations	51	Intended Harm-causation[+animate]	90	Cause oriented
attack[+human(s)]: rob: break into	2							
attack[+human(s)]: rob: make off with MONEY	1							
attack[+human(s)]: rob: hold up	1	3						
attack[+human(s)]: rob: threaten	2							
attack[+human(s)]	23	23	42	Life-threatening by human				
attack[+human(s)]: kill	1	1						
attack[+human(s)]: assault	9	10						
attack[+human(s)]: assault: raid	1							
attack[+human(s)]: assault: shoot	3	5						
attack[+human(s)]: assault: shoot, wound	1							
attack[+human(s)]: assault: shoot; rob	1							
attack[+human(s)]: assault: stab	3	3						
attack[-human(s),+animal(s)]	7	8	9	Life-threatening by nonhuman				
attack[-human(s),+animal(s)]: kill	1							
attack[-human(s),?animal]: assault[+metaphoric?]: turn on	1	1						
hit,strike: hit	3	8	18	Natural disasters	39	Disasters = Harm-causation[-animate]		
hit,strike: rock	1							
hit,strike: strike	2							
hit,strike: pound	2							
hit,strike: destroy: wreak on	1	2						
hit,strike: destroy: ravage	1							
hit,strike: roar through	1	2						
hit,strike: sweep through	1							
hit,strike: wrought devastation	1	6						
hit,strike: IMPLICIT in: earthquake	2							
hit,strike: IMPLICIT in: in PLACE	2							
hit,strike: there is	1							
hit,strike[+metaphoric, +human(s)?]: occur[=attack]	1	2	21	Social disasters[+metaphoric]				
hit,strike[+metaphoric]: hurt	1							
hit,strike[+metaphoric?]: hit	2	9						
hit,strike[+metaphoric]: hit	5							
hit,strike[+metaphoric]: paralyze	1							
hit,strike[+metaphoric]: IMPLICIT in: shocks from	1							
hit,strike[+metaphoric]: overtake	1	4						
hit,strike[+metaphoric]: take a toll	1							
hit,strike[+metaphoric]: besiege	1							
hit,strike[+metaphoric]: engulf	1							
hit,strike[+metaphoric]: occur	2	4						
hit,strike[+metaphoric]: fall on	1							
hit,strike[+metaphoric]: IMPLICIT in: in PLACE	1							
hit,strike[+metaphoric]: IMPLICIT in: problems	1	2						
hit,strike[+metaphoric]: IMPLICIT in: turmoil	1							
suffer	3	5	10	Sufferings	10	Sufferings = Harm-experience	10	Effect oriented
suffer: IMPLICIT in: victim	1							
suffer: be injured	1							
suffer: feel pain	1	3						
suffer: bring sorrow to people	1							
suffer: feel anxiety	1							
suffer: seized with	1	4						
suffer: suddenly begin a SYMPTOM	1							
suffer: experience attack[-human(s), +metaphoric]	2							

Figure 1: English translations of *osou*: variables like L1 and L2 refer to the "granularity" levels defined in Figure 2.

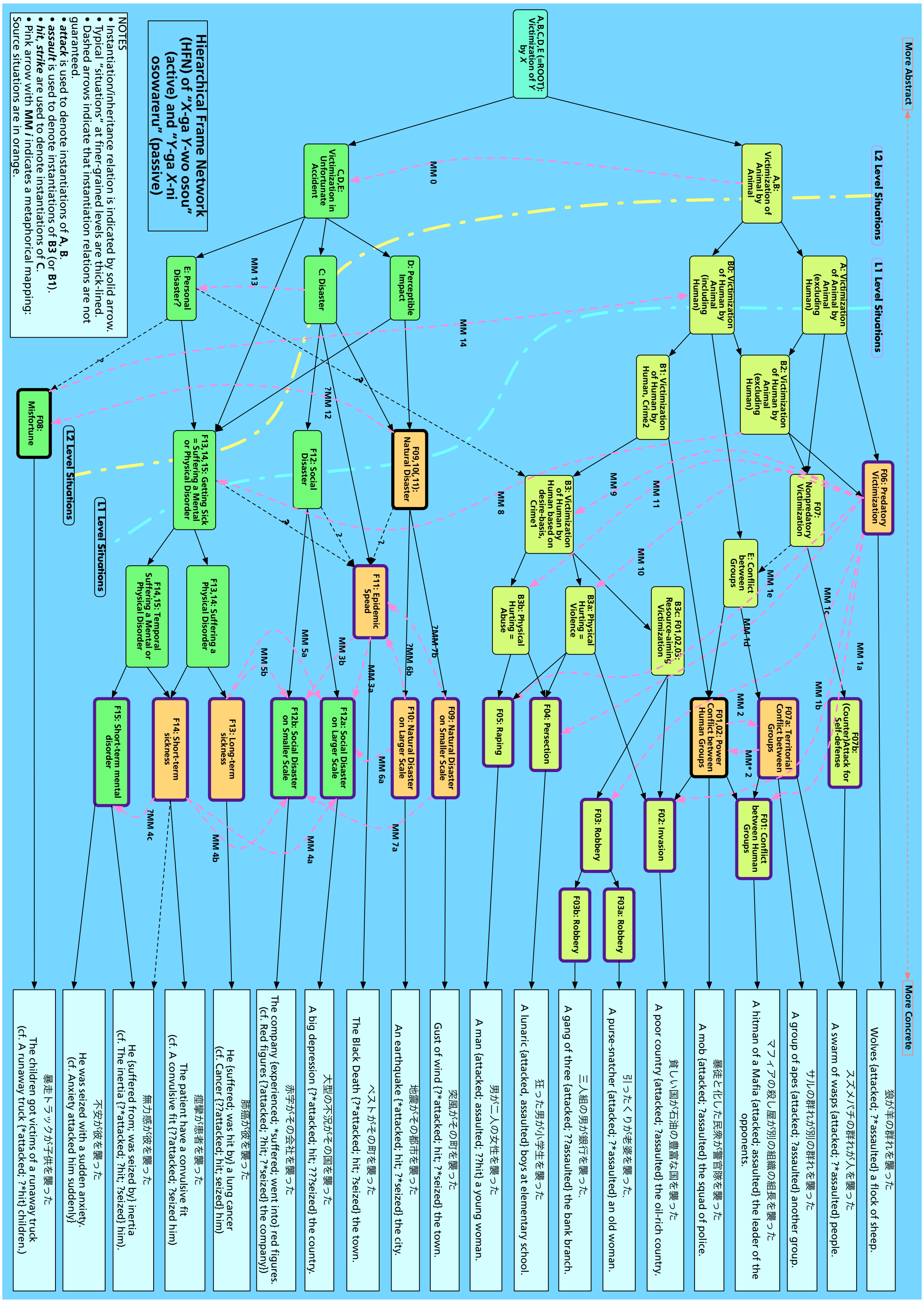

Figure 2: Lattice of situations for which an *osou*-sentence is interpreted. Solid black links indicate elaboration relationships. This lattice was designed to capture generalization from more concrete situation types (F*i* with thick border, 15 in total) at the bottom to the most abstract and generic type of situation (ROOT) at the top.

(8) **Assumption:** For every sentence $s = u_1 \cdot u_2 \cdots u_n$ with units $\{ u_1, u_2, \ldots, u_n \}$, there exists a set of semantic features $F(n) = \{ f_1, f_2, \ldots, f_n \}$ that, if chosen carefully, either characterizes the meaning of u, or, at least, differentiates the meaning of u from the meaning of another unit u' ($u \neq u'$), as far as **(i)** features are allowed to take continuous values (say, between 0 and 1.0) and are sufficiently fine-grained and **(ii)** the number of features n is large enough (and not too large).

(9) **Procedure:** Given a set of sentences $s_1 = \cdots u_1 \cdots, \ldots, s_k = \cdots u_k \cdots$ (e.g., *A family was hit by a runaway truck*) in which target units $u_1, \ldots, u_k$ (e.g., *a family*) are to be rated for semantic features $F(n) = f_1, \ldots, f_n$, ask a group of participants to rate u_i in the context of s_i against all features of $F(n)$. Average their ratings.

Under (8), the simple procedure in (9) is expected to give a good approximation to the **meaning of** u **in the specific context of** s, which should have undergone co-composition [17], rather than giving the **lexical** meaning of u.

What the procedure in (9) gives us is an approximation to the **location of** u **in a high-dimensional space** defined by the feature set $F(n)$. If $F(n)$ receives a good degree of dimensional reduction to become $F(m)$ ($n \ll m$), it is very likely that we will get a set of minimal factors $F(m)$ that account for the semantics of u.

The approach we took to represent the sentential meanings could be a "semantic vector space" approach to sentential meanings in the following senses.

First, the method we called "semantic feature rating" (SFR) is a natural extension of Osgood's **semantic differentiation method** (SD method) [3]. The differences are as follows. In the SD method, the target is **lexical** meanings, whereas in our method, the target is sentential, complex meanings. In SD, the domain of measurement is limited to a small set of usually emotional or evaluative adjectives: measurements are made on the scales of antonymous adjectives such as *good–bad*, *tall–small*, whereas in our model, the domain of measurement is general and basically open-ended propositions such as [alive(x)]: measurements are made on scales encoded by semantic features.

Admittedly, empirical research on how such features are discovered needs to be done. We turn to this in the next section.

2.3.1 Constructing a vector space through SFR

(10) Construct an HFN with a good coverage. Note: an HFN is not verb-specific: it applies to a **set** of verbs. It is not obvious, however, what verbs belong to what HFN. Determining this requires empirical research.

(11) Find a set of features such as f_1 = [visible(x)], f_2 = [carnivorous(x)] that, in combination, account for the entire HFN. This gives a feature set $F(n) = \{ f_1, f_2, \ldots, f_n \}$.

(12) Reduce $F(n)$ to $F(m)_0 = \{ f_1, f_2, \ldots, f_m \}$ ($m < n$) by removing redundant features. Such features can be detected if multivariate analysis such as Factor Analysis is applied to F.

(13) Construct the base of the semantic vector $V_0 = [f_1, f_1, \ldots, f_m]$ based on F_0. Note that V_0 defines a "semantic feature space" S specific for an HFN.

(14) Execute the SFR task for a set of sentences $\{ s_1, s_2, \ldots, s_k \}$. For each sentence, we get a semantic vector $V_j = [v_1, v_1, \ldots, v_m]$, where j denotes the index of s ($1 \leq j \leq k$) and v_i denotes the average value for feature f_i. Note that V_j defines the "location" of s_j in the high-dimensional semantic space S.

Several caveats are necessary. First, it is usually efficient if several features are determined in step (10) in the sense that [because?] they contribute to the differentiation of nodes of the HFN.

Different HFNs define different types of [values of?] V_0. Thus, V_0 needs to be modified or sometimes constructed from scratch when a different HFN is investigated. Step (12) can be omitted when F is not very big. (In fact, Step (12) was skipped in this experiment.)

Feature representation of lexical meanings is very common both in psychological research and connectionist modeling. However, as far as we know, there has not been any serious research into what features are needed for what kind of task on a large scale. A noticeable exception is [13], in which the authors attempted a "standardization" of semantic features commonly used in psychological research. If an array of semantic features is interpreted as a semantic **vector**, then a semantic space approach is possible. Many behavioral studies take this line. On the other hand, though, many psychologists seem to be skeptical of whether sentential meanings can be represented in the same way. This would explain why we did not find any previous research that attempted to represent sentential meanings in semantic vectors.

2.3.2 Why not a binary feature system?

It is traditional in linguistics to represent word meanings as "bundles of features." Our method deviates from this in that it uses continuous values for features. There are two strong reasons why we did this instead of using the traditional binary values, i.e., 1 (true) or 0 (false). First, most, if not all, features have degrees, so binary decision is simply unnatural. Second, participants are more ready to make semantic judgments with a range of confidence, expressed on a scale from very true (1) to very false (0).

Additionally, binary representation is known to have limitations and to need refinement, especially in a behavioral research setting and in computational modeling of cognitive activities, one of which is language modeling. If the value of a feature is not continuous, then the behavior of the system becomes too brittle and clumsy.

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を襲ったのは仕方のないことだった.
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/個体の規模を越える.
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.

Figure 3: 24 features/characteristics used for rating experiment.

2.3.3 Comparison with LSA

A comparison with a relatively well-known model *Latent Semantic Analysis* (LSA) [12] may be helpful here. In LSA, unlike our approach based on behavioral data, semantic vectors are constructed from corpus data (through a dimension-reduction technique called "singular value decomposition"), though the idea of representing word meanings in vector form is shared. The main difference is that semantic vectors in LSA are very big and that the notion of features is no longer tenable.

Furthermore, sentential meanings are constructed under strict compositionality: the meaning of a sentence $s = w_1 \cdot w_2 \cdots w_n$ is defined as the (logical) conjunction of semantic vectors $v_1, v_2, \ldots, v_n$ (where v_i denotes the semantic vector of word w_i), which corresponds to a particular point in a semantic space.

3 Experiments

3.1 Procedure

In our SFR task for *osou*-sentences (explained in §3.2), participants are presented with Japanese sentences in which (i) the main verb is *osou* (active form) or *osowareru* (passive form) and (ii) either the subject or object NP is a bisyllabic nonce word, which, therefore, has no lexical meaning.

Participants were asked to rate each of the 24 features in Figure 3 on a five-point scale (from "very true" 5.0 to "very false" 1). The results were averaged. Several types of multivariate analyses (e.g., Principal Component Analysis (PCA) and Factor Analysis (FA)) were applied to it.

3.2 Materials

A Japanese sentence of the form "X-ga Y-wo osou" [active] or "Y-ga X-ni osowareru" [passive]) denotes a situation in which Y is victimized by X, a ⟨harm-causer⟩.

All Japanese examples (in the 6th column) of Figure 4 were constructed for a lattice of situations, presented in Figure 2, with 15 lowermost, most finely grained levels (F01, . . . , F15).[6] The lattice of situations in Figure 2 was constructed from a frame-based manual analysis of the 413 examples from a corpus [20], whose validity was confirmed by an independent psychological experiment [15]. We assumed that contextual effects on u_i from its context s_i were factored out and controlled in this way, though this point could, we are aware, be controversial.

4 Results and Discussion

4.1 Main results

For both *osou*- and *osware*-sentences, nonce words were rated very much like real words in the way pre-

6) This does not mean, however, that there are no classifications with a finer granularity than F01, . . . , F15. We plan to conduct an experiment to see if hyper-fine-grained situations in which ⟨attack by a large predatory animal⟩ and ⟨attack by a small predatory animal⟩ can show as much convergence as we had in our experiment that confirmed those 15 situations.

Situation ID	Subject NP denoting Victim	PP denoting Harm-causer	Transliteration (word-by-word translation from Japanese assuming that *osou* translates to *attack*)	Natural Translation	Original Example (in Passive form)
F01: Power conflict between human groups	The President	an assassin	The President was attacked by an assassin.	The President was assaulted by an assasin.	大統領が 暗殺者に 襲われた。
F02: Invasion	a country	an(other) armed country	A country was attacked by another armed country.	A country was attacked by another armed country.	ある国が 軍事国に 襲われた。
F03: Robbery on larger scales	a bank branch	a masked man	A bank branch was attacked by a masked man.	A bank branch was attacked by a masked man.	銀行が 覆面の男に 襲われた。
F03: Robbery on smaller scales	an old lady	a purse snatcher	An old lady was attacked by a purse snatcher.	An old lady was assaulted by a purse snatcher.	ある老婆が ひったくりに 襲われた。
F04: Persecution, Violence	passengers-by	a lunatic man	Passengers-by were attacked by a lunatic man.	Passengers-by were assaulted by a lunatic man.	通行人が 精神障害の男に 襲われた。
F05: Sexual assault	a woman	a pervert	A woman was attacked by a pervert.	A woman was sexually assaulted a pervert.	ある女性が 性的倒錯者に 襲われた。
F06: Preying animal attack; Predation	zebras	lions	Zebras were attacked by lions.	Zebras were attacked by lions.	シマウマが ライオンに 襲われた。
F07: Nonpreying animal attack, usually for defence	children	wasps	Children were attacked by wasps.	Children were attacked by wasps.	子どもが スズメバチに 襲われた。
F08: Accident	a family	a runaway truck	A family was attacked by a runaway truck.	A family got the victim of an accident by a runaway truck.	ある家族が 暴走トラックに 襲われた。
F09: Natural disater on smaller scales	a town	gust of wind	A town was attacked by gust of wind.	A town was hit by gust of wind.	ある集落が 突風に 襲われた。
F10: Natural disaster on larger scales	a local area	a hurricane, a typhoon	An area was attacked by a hurricane.	An area was hit by a hurricane.	ある地方が 台風に 襲われた。
F11: Epidemic spread	a city	influenza, Black Death	A city was attacked by influenza.	A city was hit by influenza.	ある都市が インフルエンザの流行に 襲われた。
F12: Social disaster	the stock market	a debacle (or a downturn, sharp fall)	The stock market was attacked by a sharp fall.	The stock market was hit by a sharp fall.	株式市場が 株価の暴落に 襲われた。
F13: Long-term sickness	a man	cancer, malignant tumor	A man was attacked by cancer.	A man was seized by cancer; A man suffered cancer.	ある人が 悪性のガンに 襲われた。
F14,15: Short-term mental disorder OR sickness	an old man	panic	An old man was attacked by panic.	An old man was seized by panic.	ある老人が 不安に 襲われた。
F14: Short-term sickness/symptom	a man	a sharp pain	A man was attacked by a sharp pain.	A man was seized by a sharp pain; A man suffered a sharp pain.	ある男性が 激痛に 襲われた。
F15: Short-term mental disorder	a young man	strong jealousy	A young man was attacked by a strong jealousy	A young man was seized by a strong jealousy	ある若者が 激しい嫉妬に 襲われた。
NONSENSICAL	zebras	a masked man	Zebras were attacked by a masked man.	Zebras were attacked by a masked man.	シマウマが 覆面の男に 襲われた。

Figure 4: Materials used for experiment, with English translations (passive cases only).

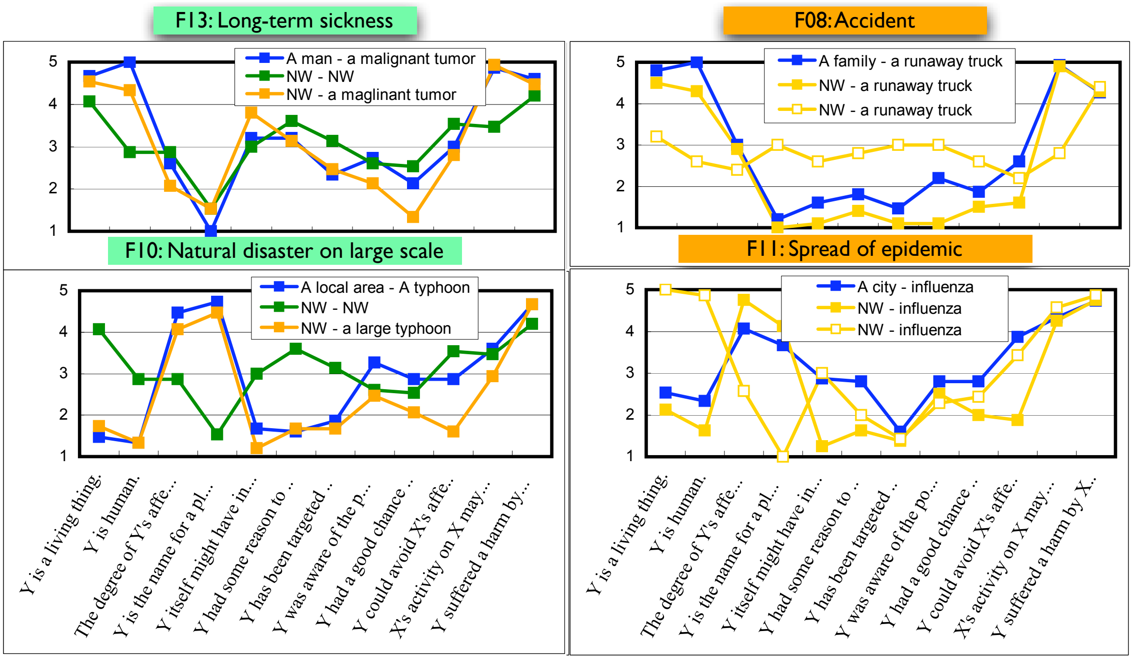

Figure 5: SFR profiles for F13 and F10 (Group 1) and F08 and F11 (Group 2) for "⟨ victim(Y) ⟩-ga ⟨ harm-causer(X) ⟩-ni osowareta" (X= NW (nonce word), Y=RW (real word)) contexts, comparing the results for the real-word rating (in blue) and two response patterns. For Group 2, two response patterns (both in orange) were differentiated. For comparison, NW-NW response (Baseline 2) is marked in green for F13 and F10.

dicted by the A-to-S effect. This is indicated by the fact that in cases F13 and F10, for example, nonce words for ⟨ victim ⟩s were feature-rated as like real words (*a man* and *a local area*, respectively) in the corresponding full sentences, as indicated by the left-side profiles in Figure 5. The same was true of the nonce words for the ⟨ harm-causer ⟩.

Green graphs for F13 and F10 indicate the NW-NW response. It is reasonable to think, from theory, that these represent the *lexical* meaning of *oso(ware)u*. Differences from it indicate the effects of co-composition. It is also reasonable to think of the specific situation closest to the NW-NW response as being the prototypical situation of the events that can be referred to by *oso(warer)u*. It turned out that F06: ⟨ predatory victimization ⟩ was the nearest situation when the Euclidean distance in the space defined by the first three principal components of the semantic features was used as a measure of dissimilarity. We found that this was a reasonable result.

4.2 Discussion

4.2.1 When metonymic adjustment is called for

SFR patterns for F08 and F11 behaved somewhat differently from the others. Unlike other cases, they reflect "logical metonymy" for ⟨ victim ⟩, giving rise to two different rating patterns, which correspond to two different orange graphs overlaid in the right-side profiles in Figure 5. It seems that in one response pattern, nonce words for ⟨ victim ⟩ were characterized like [individuals] (e.g., *a family*) that were at that location at that time;[7)] in another, they were characterized like ⟨ location ⟩s where ⟨ victims ⟩ were located at the time of victimization. For the former case, a metonymically based reference shift from [place] to [individuals] is observed.

Despite problems like this, it was found that the interpretations of *oso(warer)u*-sentences do not exceed the range of possible interpretations specified by the situation lattice in Figure 2. This suggests that conventional metaphors are lexicalized and are not as productive as claimed by [11], in agreement with the claim in [1].

4.2.2 Strengths of A-to-S effects

Different nouns have different degrees of A-to-S effects. This is not at all surprising, but it should be noted that nouns showing strong A-to-S effects are names for **representative instances** of ⟨ harm-causer ⟩ or ⟨ victim ⟩. Nouns that denote ⟨ harm-causer ⟩s, on average, were found to have stronger A-to-S effects than nouns that denote ⟨ victim ⟩ when combined with *osou*.

4.2.3 Scalability

There is no reason to doubt that this result can be extended to other constructions because there seems to be nothing special about the behavior of the *oso(ware)ru* construction that we investigated. It shows the normal behavior of a polysemous verb.

It is not at all easy, however, to see what will actually happen with other constructions. First of all, if we decided to do the same experiment with another verb V, a different set of situations, desirably in the form of an HFN, would need to be constructed for V. For example, if we decided to test *y-ga x-kara nigeru* (meaning "*y* ran away from *x*" in English), we would need to construct an HFN for this verb. This task would, admittedly, be very painstaking. We have a hope of semi-automated of this take using the method tested in [9].

4.2.4 Trouble with the basic units of evocation

It is hard to reconcile what we have shown with the traditional account of semantic interpretation in which meaning construction is define as a rule-governed, compositional process, but let us try.

As Frame Semantics tells us, words and phrases evoke specific situations, or "frames" in the sense of FS/BFN. Evoked frames are integrated, thereby giving the semantic representation of a complex unit, say of a sentence. Frame integration should be a co-compositional process. Nothing is wrong so far. But here comes an annoying question, What are the "basic units" of situation/frame evocation? — Are they words or larger units like collocational patterns?

Our results strongly suggest the that **larger units have stronger situation/frame-evocation effects than smaller units like words**. This implies that collocations patterns are better units of evocation than words. This poses a challange to any lexicon-building attempt because it questions one of its most important assumptions: Is it really promising to try to build a lexicon that should provide a (desirably) necessary and sufficient information for semantic description? because a lexicon, by its defintion, mainly, if not only, gathers meanings of lexical items, typically words. If not, what shall we do?

This also implies that the semantics of regular units may not be as much compositional as is usually supposed to be, because if pairings of surface forms and their meanings are evocation-based, their semantics **need not** be compositional in the sense of traditional linguistics and logic. In this scenario, semantic specifications are directly associated with collocational patterns, and it is very likely that what Hunston and Francis [8] call "patterns" and what Wray [21] calls "formulaic language" play more vital role than in the traditional account. A similar insight plays an important role in *Corpus Pattern Analysis* (CPA) [7, 18].

Thus, there seems to be a serious need to establish a good theory of the semantics of collocational patterns, i.e., "superlexical" units, on the one hand, and to reconcile between the description of lexical items and that of collocational patterns, on the other. We are not profesional lexicographers, and clearly are not qualified to propose any solution to this problem, but we can suggest, we believe, some workaround that would make a lexicon-building task more realistic, building on the seminal work by [5]. We also hope it is compatible with the basic idea underlying CPA.

[7)] For F08, *a family* is understood to refer to its members as individuals.

The workaround we have in mind is to make a clever use of ontological information specified in HFNs like the one in Figure 2. Let us explain the basic idea with a few examples. Suppose we are on specifying the semantics of “NP *attack* NP.” Sentences like *The lions attacked a herd of impalas*, *A group of killer whales attacked a humpback whale* can be seen as instances of the [⟨ predator ⟩ *attack* ⟨ prey ⟩] schema interpreted against the situation of ⟨predation⟩. Assuming a proper defintion of ⟨predation⟩, we can say [*the lions* instance-of ⟨predator⟩], [*(a herd of) impalas* instance-of ⟨prey⟩], [*killer whales* instance-of ⟨predator⟩], and [*a humpback whale* instance-of ⟨ prey ⟩], using instance-of link. Recall that the HFN in Figure 2 is the partical specification of the ontology of harm-causation, of which ⟨ predation ⟩ is an instance.

It is important to note that many, if not all, situations are associated with **role names** equivalent to *predator* and *prey* for the ⟨ predation ⟩ situation. For one, *victim* is the role-denoting noun that is valid to denote any instance of *y* on the entire HFN in Figure 2. We have role hierarchies like [⟨ bank ⟩ is-a ⟨ victim ⟩] for ⟨ bank robbery ⟩, [⟨prey⟩ is-a ⟨victim⟩] for ⟨predation⟩. Based on this, we suggest that it is promising to build a lexicon in which senses of the arguments of a predicate (e.g., *attack*) are defined by referring to the **role hierarchies** derived form the situation hierarchies like the HFN in Figure 2. A clever use of this kind of information should make a lexicon more realistic and amenable to the information encoded by collocational patterns, we believe.

5 Conclusion

In this paper, we tried, by presenting psychological evidence, to argue for a radically selectional theory of semantic interpretation based on Frame Semantics, and contrasted it with purely constructivist theories of semantic interpretation. We also pointed out that logically based models of sentential meaning need to be somehow modified to make them compatible with vector space models of it, because there is a large gap between the logical forms and the behavioral data obtained through psychological experiments. We suggested that lexicon-building task can be made more realistic if we employ roles hierarchies derived from situation hierarchies.

References

[1] N. Asher and A. Lascarides (2001). Metaphor in discourse. In P. Bouillon and F. Busa, eds., *The Language of Word Meaning*, pp. 262–290. Cambridge University Press.

[2] Reinhard Blutner and Henk Zeevat, eds. (2004) *Optimality Theory and Pragmatics*. Palgrave Macmillan, Houndmills, Basingstoke, Hampshire.

[3] Osgood C. E. (1952). The nature and measurement of meaning. *Psychological Bulletin*, 49(3):197–237.

[4] C. J. Fillmore (1985). Frames and the semantics of understanding. *Quaderni di Semantica*, 6(2): 222–254.

[5] C. J. Fillmore and B. T. S. Atkins (1994). Starting where the dictionaries stop: The challenge for computational lexicography. In B. T. S. Atkins and A. Zampoli, eds., *Compuational Approaches to the Lexicon*, pp. 349–393. Clarendon Press, Oxford.

[6] T. Fontenelle, ed. (2003). *FrameNet and Frame Semantics*: A Special Issue of *International Journal of Lexicography*, 16 (3).

[7] P. Hanks and J. Pustejovsky (2005). A pattern dictionary for natural language processing. *Revue française de la linguistique appliquée*, 10 (2).

[8] S. Hunston and G. Francis (2000) *Pattern Grammar: A Corpus-Driven Approach to the Lexical Grammar of English*. Amsterdam: John Benjamins.

[9] Kanamaru, T., M. Murata, K. Kuroda, and H. Isahara (2005). Obtaining Japanese lexical units for semantic frames from Berkeley FrameNet using a bilingual corpus. In *Proceedings of 6th International Workshop on Linguistically Interpreted Corpora*, pp. 11–20.

[10] E. Kako (2006). The semantics of syntactic frames. *Language and Cognitive Processes*, 21(5): 562–575.

[11] G. Lakoff and M. Johnson (1980). *Metaphors We Live By*. University of Chicago Press.

[12] T. K. Landauer and S. T. Dumais (1997). A solution to Plato's problem: The Latent Semantic Analysis theory of the acquisition, induction, and representation of knowledge. *Psychological Review*, 104: 211–240.

[13] K. McRae, G. Cree, M. S. Seidenberg, and Ch. McNorgan (2005). Semantic feature production norms for a large set of living and nonliving things. *Behavior Research Methods, Instruments, & Computers*, 37(4): 547–559.

[14] K. Nakamoto and K. Kuroda (in press) Representing selectional restrictions in terms of semantic frames equated with situational schemas: A case study of the japanese verb *osou*. *Studies in Language Science*, 7.

[15] K. Nakamoto, K. Kuroda, and H. Nozawa (2005) Proposing the feature rating task as a(nother) powerful method to explore sentence meanings. *Japanese Journal of Cognitive Psychology*, 3 (1): 65–81. (written in Japanese).

[16] A. S. Prince and P. Smolensky (1993). *Optimality Theory: Constraint Interaction in Generative Grammar*. Technical Report No. 2, Rutgers University Center for Cognitive Science.

[17] J. Pustejovsky (1995). *The Generative Lexicon*. MIT Press.

[18] A. Rumshisky and J. Pustejovsky (2006). Inducing sense-discriminating context patterns from sense-tagged corpora. In *Proceedings of LREC 2006, Genoa, Italy*.

[19] J. Ruppenhofer, M. Ellsworth, M. R. L. Petruck, and C. Johnson (2005). *FrameNet: Theory and Practice*. [http://framenet.icsi.berkeley.edu/book/book.html]

[20] M. Utiyama and H. Isahara (2003). Reliable measures for aligning Japanese-English news articles and sentences. In *ACL-2003*, pp. 72–79.

[21] A. Wray (2002). *Formulaic Language and the Lexicon*. Cambridge University Press, Cambridge/New York.