The Correct Place of Lexical Semantics in Interlingual MT

Lori LEVIN and Sergei NIRENBURG
Center for Machine Translation
Carnegie Mellon University
Pittsburgh, PA 15213, U.S.A.

1. Introduction

Interlingual MT has typically come to include a syntactic analysis of source language (SL) text followed by its semantic interpretation and representation in terms of a text meaning representation (TMR) scheme, an interlingua. Recently two distinct views of the nature of the interlingua have become current – one based on a world model (e.g., Nirenburg et al., 1992) and another one based on the notion of lexical conceptual structure (LCS) (e.g., Dorr, 1992). In this paper we analyze the role of LCS in the extraction of text meaning and argue that, though it cannot be considered an interlingua when used by itself, it contributes significantly to the specification of an actual interlingua. The task of an interlingual MT system builder is, then, to find a way to integrate the information provided in LCS into an ontology-motivated text meaning representation serving as interlingua. In this paper, we propose a model for this integration and illustrate the processes and static knowledge sources involved, centrally including the lexicon.

In Section 2 we propose a model of MT that involves both an LCS-based lexical semantic structure and a TMR that is not based on LCS. Because our lexicon formalism does not represent LCSs, but semantic role names that serve as labels for LCS variables, we will use the abbreviation SDLS (for syntax-driven lexical semantics, Nirenburg and Levin, 1992) in reference to our system instead of LCS. We argue that TMR and SDLS are both necessary and that they are distinct. This model forms the basis of lexical-semantic treatment of texts in the multilingual MT project *Mikrokosmos*. In Section 3 we present specific examples as analyzed in Mikrokosmos. We illustrate the static knowledge sources (primarily the lexicon) and the representations that are produced (syntax, lexical semantics, and TMR). The Mikrokosmos model is based on a theory of form-to-meaning correspondence which relies on the concept of a society of microtheories integrated in a non-stratificational manner. We briefly sketch the main points of this theory in the final section of this paper.

2. The model

Traditionally, interlingual MT systems which employ a full-blown syntactic module (e.g., KBMT-89 (Goodman and Nirenburg, 1992) or KANT (Carbonell et al., 1992)) use a single mapping between syntactic structure and interlingua. In Mikrokosmos, we propose a different model, as illustrated in Figure 1. Lexical-conceptual structures (LCSs) have been suggested as meaning representations for natural language sentences produced in accordance with the semantic theory developed by Hale and Jackendoff (e.g., Jackendoff, 1983) and used in MT-related experiments by Dorr (Dorr, 1993). The interlingua text (or text meaning representation, TMR) is a structure which represents meaning of texts in accordance with the ontology-oriented approach to computational semantics (see Nirenburg and Levin, 1992).

It is convenient to structure our argument for this model around the questions below (referring to labels in Figure 1), which we will discuss one-by-one in the following subs

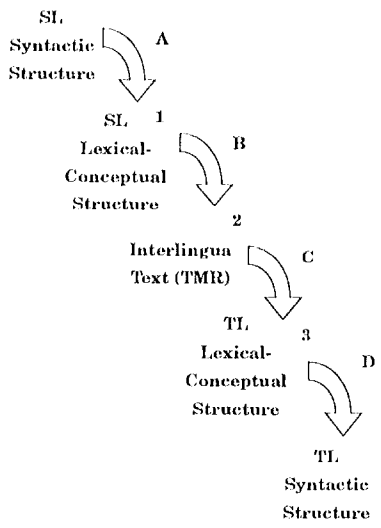

Figure 1. *Data Flow in a KBMT System with 2 Semantic Modules.*

1. How similar are structures 1 and 3? How are they different?
2. How similar are mappings A and D? How are they different?
3. How is structure 2 different from structures 1 and 3?
4. Why are representations 1, 2, and 3 all necessary?

2.1. Are Lexical-Conceptual Structures Language-Universal?

Attempts have been made to use LCSs as interlinguas for MT (notably, Dorr, 1993). The impetus for such work is provided by observations that in many cases LCSs for translation equivalents are, in fact, identical. The many cases in which LCSs are not identical across languages pose problems for this approach. Methodologically, therefore, the type of work in LCS-as-interlingua projects is finding ways of resolving each such case, based on observing cross-linguistic divergences in realizing meanings. There is a danger that some of the divergences will prove untreatable at the LCS level and, alternatively, that solutions for some problems will necessitate changes to the nature of the representation which will make the resulting structure resemble the original LCS in progressively smaller ways. The problematic cases will be those in which translation equivalents can

have different lexical semantics. We will mention two such cases here.

The first problem arises in the context of a complex event, such as a merger of two companies, which can be described by mentioning any of its parts (bids, negotiations, etc.). This is particularly problematic when different languages, by convention or for ease of expression, refer to different parts of the complex event. In fact, such divergences exist even within one language. For example, you can go to a meeting (directed motion), attend a meeting (activity), or be at a meeting (state). Similarly, while in English one *takes* a taxi, using a transitive verb, the corresponding Japanese for the same event it *takusi ni noru* (get on, board, ride in a taxi), using an intransitive verb with a goal argument. Even seemingly atomic events and states can be broken down into their aspectual components to consist of events leading up to changes of state that result in new states. For example, the situation of knowing something can be expressed in English using the stative verb *know* or in Japanese using a non-stative verb *siru* (come to know) in its resultative form *site iru* (Lit: have come to know). In examples such as these, there will be no direct correspondence at the level of lexical semantics in individual languages.

The second circumstance in which translation equivalents have different lexical semantics is that an element of meaning that is expressed as an argument-taking predicate in one language might not be expressed as an argument-taking predicate in another language. Well-known examples from MT literature include *like* vs. *gern*, *venir de* vs. *just*, etc. However, this phenomenon is much more widespread than normally acknowledged in the MT literature. Things that are expressed as main or auxiliary verbs in English, but are not verbs at all in Japanese include many high-frequency meaning elements such as phase (*begin, continue, finish*), modality (*must/should, plan, expect, try*), and evidentiality (*seem, appear, look like*). In fact, the syntactic means for encoding these types of meaning vary wildly among languages, going far beyond the well-known verb-adverb divergences. This is why in the Mikrokosmos interlingua we represent such elements of meaning as features or operators that scope over clauses and propositions.

2.2. How is an Interlingua Different from an SDLS Output?

In the cases described above in which a single event is described with different lexical semantics the meaning shared by each member in the set of paraphrases makes a better candidate for the interlingual semantic representation than does the lexical semantics; and it is this type of meaning that we are striving to extract and represent in the interlingua text in Mikrokosmos. Additionally, while SDLS concentrates on the "who-did-what-to-whom" aspect of text meaning, TMRs contain additional meaning facets, such as aspect, modality, evidentiality, speech act, reference, etc. Finally, as TMRs are not based on the lexical semantics of one particular language, there is no special benefit to be accrued from the imposition of the requirement to preserve predicate-argument structures.

2.3. Universals of Semantic Role Assignment

It is very enticing to be able to apply principles of lexical mapping theory cross-linguistically. Similarities that have been observed across languages involve linkings of semantic roles to syntactic positions or grammatical functions, transitivity alternations, and verb classes. The latter have been described in some detail for English by B. Levin (1993) and others. Thus, *to the extent that the hypothesis of cross-linguistic equivalence holds*, the description of similar phenomena in other languages, for the purposes of MT, becomes much simpler, if not utterly trivial.

However, languages, as a rule, have different transitivity alternations (Mitamura 1989) and even when they have a similar transitivity alternation, the classes of verbs to which they apply may be different. See Mahmoud 1989 for a discussion of the differences in the verbs that undergo the causative-inchoative alternation in English and Arabic.[1] It is, of course, desirable to take advantage of universals, but it is also necessary to have a system that is flexible enough to accommodate cross-linguistic variation.

2.4. Integration of SDLS into Interlingual MT

Taking a position on the necessity of both SDLS and TMR has to be based on a general approach to unraveling the form-meaning correspondence. For example, to make a TMR for *John began to read* we need to identify a number of meaning elements, primarily that something took place before the time of speech, which was the beginning phase of a reading event carried out by John.[2] How do we find these pieces of information? Time before the time of speech is indicated by the *morphology* of "began". The beginning phase is typically indicated *lexically* by the verb *begin* in English. We know that it is the beginning phase of reading because the *syntax* module tells us that *to read* is the complement of *begin*. We know that John is reading because *John* is the subject of *begin* (once again, the *syntactic* module produced this element of information), whose lexical properties tell us that *John* is also understood as the subject of the complement clause. In other words, it is the predicate argument structure of *begin* (produced by the syntax-to-SDLS mapping procedure in the *lexicon* entry for *begin*) that tells us where to find many of the relevant pieces of information.

Having thus served the purpose of identifying a part of the semantic dependency to be represented in the final TMR (just as the findings of other system modules played their assigned roles as clues for determining parts of the TMR structure), the predicate argument stucture can then be discarded. In the following section we give some detailed examples of the mappings involved in producing SDLS output structures and TMRs as well as relevant parts of lexicon entries.

3. Some Examples

Examples in Figures 2, 3 and 4 contain a number of representative phenomena which underscore the differences between SDLSs and TMRs as well as illustrate how the two structures co-exist in the Mikrokosmos processing model. In doing so, we also describe a lexicon design which accommodates both structures. In all three examples the SDLS is just one of the clues for determining a component of meaning, and is not preserved isomorphi-

[1] Incidentally, therefore verb classes are not suitable as semantic hierarchies for ontology (Mitamura 1989).

[2] It could also be the beginning phase of a habit of reading instead just one instance of reading — there is no way to determine which in the absence of context.

cally in the TMR. The examples also illustrate the use of constructions (Fillmore et al. 1988, Fillmore and Kay 1992) as a unit of analysis alongside words, and show that treatment of MT divergences in this approach simply falls out of the general model. The languages used for illustration are English, Russian, and Japanese. Since the system is symmetrical, we do not identify which is the source language and which is the target language in each example.

For each example, we list a TMR, which is the same for all of the languages, as well as syntactic structures, semantic role assignments (SDLS), and lexical entries for each language. It should be apparent that the TMR is not necessarily isomorphic to the SDLS of any of the languages, and that sentences from different languages can correspond to the same TMR even if their syntactic and SDLS representations are not isomorphic. The Mikrokosmos TMR structure consists of clauses which roughly correspond to the "who did what to whom" component of meaning but also includes such components as speech acts, speaker attitudes, indices of the speech situation, stylistic factors as well as relations (e.g., temporal ones) among any of the above, and other elements.

The lexical entries include three zones—syntax, semantic role assignment, and mapping to TMR. (The first and third zones are discussed by Meyer et al. 1991.) The first zone specifies an LFG-style (Bresnan 1982) syntactic subcategorization frame of a predicate, including which grammatical functions (subject, object, complement, etc.) the predicate must appear with and any requirements the predicate has of those functions (case, syntactic category, specific lexical items, etc.). The second zone, also in the spirit of LFG, specifies a mapping between the grammatical functions governed by a predicate and the semantic roles it assigns. Semantic role assignment is indicated by coindexing of a syntactic slot and a semantic role slot. The semantic role names used in the examples are simply labels for argument positions in lexical conceptual structures, which are not shown here. The syntax and semantic role assignment zones serve the purpose of locating the important participants in the sentence. For example, they might tell us that the experiencer argument is in the subject slot with dative case, or that the phrase functioning as the theme argument is found in the object position. They are also important in capturing both language-specific generalizations about verb classes and universals of semantic role assignment. For these reasons, the syntax and semantic role zones are crucial, and therefore must be included even in cases in which they differ drastically from the TMR.

The third zone of the lexical entry specifies portion of TMR that is associated with a lexical item and how the components of the TMR correspond to the components of the syntactic and semantic role zones. We have chosen examples in which the TMR is not isomorphic to the syntactic and lexical semantic zones. In most of the examples, a lexical item specifies that one of its complements heads the associated TMR. In these cases, the syntactic head of the sentence corresponds to some kind of scope-taking operator or a simple feature-value pair in TMR.

The examples, incidentally, illustrate our treatment of MT divergences—situations in which a source language sentence and its target language translation differ significantly in syntactic structure, syntactic category, or predicate-argument structure. No special mechanisms are needed to treat MT divergences in our model. All that is needed in order to translate a sentence involving a divergence are source and target language lexical entries of the sort illustrated here that map different syntactic structures onto the same TMR. The representations and mechanisms shown in the lexical entries are motivated for non-divergent examples and do not have to be modified to deal with divergent examples. This is because source and target language sentences are not normally expected to be isomorphic to the TMR or to each other.

Another important feature of our model is that it considers *constructions* to be basic lexical units along with words. Following Fillmore et al., 1988, we define constructions as (possibly, discontiguous) syntactic structure or productive syntactic pattern whose meaning it is often impossible to derive solely based on the meanings of its components. In other words, a construction is a combination of a syntactic structure and the associated semantic and pragmatic representations which, once detected, do not have to be compositionally produced by a TMR extractor. Constructions are typically ways of expressing a meaning that are conventional in the sense that they are frozen, and not synchronically derivable from general principles, even if they once were. Note that a formalism such as the HPSG-like sign or the dictionary structure of the ACQUILEX project can be made to support such an idea, as Fillmore and Kay (1992) show.

4. Lexical Semantics in an Overall Theory of Form-Meaning Correspondence

The Mikrokosmos project is based on a theory of form-meaning correspondence, whose underlying assumptions can be stated as follows:

- Meanings are extracted from texts on the basis of all and any available clues (e.g., syntactic, morphological, and lexical properties of an utterance). The extraction of meaning consists of constructing the most plausible, though usually defeasible, hypothesis that is compatible with the evidence, making it an abductive process (Hobbs, 1991).
- The processing of clues in Mikrokosmos is grouped into microtheories for elements of meaning such as predicate-argument relations, aspect, temporal relations, modality, evidentiality, etc. Each microtheory specifies the ways to construct TMRs for some aspect of meaning by identifying the various syntactic, morphological, and lexical clues for that element of meaning in individual languages.
- In integrating the microtheories, Mikrokosmos rejects the pure stratificational approach shared by such otherwise diverse models as AI NLP semantics (e.g., Hirst, 1987) or Mel'čuk's MTM (e.g., Mel'čuk, 1981). Knowledge from all kinds of areas coexists in the same rules for the determination of meaning units.
- The clues (pieces of evidence) for an element of meaning can interact in complex ways. Clues can reinforce or contradict each other. Coercion is possible in situations in which the clues conflict. Interpretation of a clue can be dependent on which other clues are present.

Figure 3. Speech Act

English: *Can you buy me a book?*
Japanese: *Hon o katte moraemasen ka?*

TMR

clause-1
- head: buy-1
- agent: *hearer*
- theme: book-1
- aspect:
  - phase: none[1]
  - duration: momentary
  - iteration: single

speech-act-1
- type: request-action
- scope: clause-1
- speaker: *speaker*
- hearer: *hearer*

relation-1
- type: temporal-before
- from: time-of-speech
- to: time-of(clause-1)

relation-2
- type: possession
- from: *speaker*
- to: book-1

relation-3
- type: temporal-before
- from: "time of clause-1"
- to: "time of relation-2"

	English		Japanese
Syntactic Structure	PREDICATE can SUBJECT you COMPLEMENT PREDICATE buy SUBJECT you OBJECT me OBJECT2 book		PREDICATE: morau (receive) MOOD: potential interrogative TENSE: non-past SUBJECT: speaker OBJECT: hon (book) OBJECT2: hearer COMPLEMENT: PREDICATE: kau (buy) SUBJECT: hearer OBJECT: hon (book)[4]
Semantic Role Assignment	CAN proposition: buy agent:: you beneficiary: me theme: book		MORAU (receive) recipient: speaker source: hearer favor: kau (buy) agent: hearer theme: hon (book)
Lexicon: Syntactic Structure	CAN[2] predicate: [0] can subject: [1] root: you complement: [2] subject: [1]	BUY predicate: [0] buy subject: [1] object: [2] object-2: [3]	MORAU[5] predicate: [0] tense: non-past mood: potential subject: [1] root: speaker object: [2] object-2: [3] root: hearer complement: [4] inflection: gerund subject: [3] object: [2]
Lexicon: Semantic Role Assignment	proposition: [2]	agent: [1] beneficiary: [2] theme: [3]	recipient: [1] source: [3] favor: [4]
Lexicon: Mapping to TMR	clause [3] head: meaning-of ([2]) speech-act [4] type: request-action scope: [3] speaker: *speaker* hearer: *hearer* relation [5] type: temporal-before from: time-of-speech to: time-of ([3])	clause [4] head: *BUY agent: meaning-of ([1]) theme: meaning-of ([3]) relation [5] type: possession from: meaning-of ([2]) to: meaning-of ([3])[3] relation [6] type: temporal-before from: time-of ([4]) to: time-of ([5])	clause [5] head: meaning-of ([4]) speech-act [6] type: request-action scope: [5] speaker: *speaker* hearer: *hearer* relation [7] type: temporal-before from: time-of-speech to: time-of ([5])

[1] When a reference is made to an entire event, the phase is none of *begin*, *end* or *continue*.
[2] This lexical entry is for the construction "Can you X," which we take to be a conventional way of making a request. There are other lexical entries for other senses of *can*.
[3] Note that this relation (that the beneficiary will possess the book) is not explicit in the SDLS, but it is part of the default meaning of the sentence. The Mikrokosmos architecture allows for this default to be overridden in contexts that contain conflicting information.
[4] The constitutent structure of this sentence would be mono-clausal, but the functional structure as shown here is bi-clausal. It is possible that it should even be tri-clausal depending on the analysis of the potential morpheme. See Matsumoto (1992) for a discussion of these issues.
[5] This entry, like the English entry for *can* describes a construction that is conventionally used for making requests. The verb *morau* (receive) has other entries as well. The potential and negative morphemes are iptional in the request construction. It is important, though, that they simply soften the construction and do not take on their literal meanings (Horiguchi 1993).

Figure 2. Phasal Verbs

English: *John started to speak.*
Russian: *Dzhon zagovoril.*

TMR

clause-1
- head: speak-1
- agent: John
- aspect:
  - phase: begin
  - duration: prolonged
  - iteration: single[1]

	English		Russian
Syntactic Structure	PREDICATE start SUBJECT John COMPLEMENT PREDICATE speak SUBJECT John		PREDICATE: zagovorit' SUBJECT: Dzhon
Semantic Role Assignment	START theme: John goal: SPEAK agent: John		ZAGOVORIT' agent: Dzhon[2]
Lexicon: Syntactic Structure	START predicate: [0] start subject: [1] complement: [2] subject: [1] inflection: infinitive	SPEAK predicate: [0] speak subject: [1]	ZAGOVORIT' predicate: [0] govorit' subject: [1] phase: inceptive aspect: perfective
Lexicon: Semantic Role Assignment	theme: [1] goal: [2]	agent: [1]	agent: [1]
Lexicon: Mapping to TMR	clause [3] head: meaning-of ([2]) aspect: phase: begin	clause [2] head: *SPEAK agent: meaning-of([1])	clause [2] head: *SPEAK agent: meaning-of ([1]) aspect: phase: begin[3]

[1] *OR* iteration: multiple (in, e.g., "John started to speak at union meetings").
[2] We are treating the sentence as mono-clausal, even at the level of argument structure, because the sentence "Dzhon zagovoril kazhdyj den" (John started to speak every day) does not exhibit the same ambiguity (Every day John started to speak/John started a habit of speaking every day) as the bi-clausal English sentence. If this reasoning is wrong, we can easily have a bi-clausal lexical semantic structure associated with the mono-clausal syntactic structure.
[3] To the extent that morphology is productive, this entry can be derived automatically from the entry for the base form "govorit'".

Figure 4. Modality

English: *You'd better go.*
Japanese:
Itta hoo ga ii.
go-past alternative subj. good.
Lit: The alternative that you went is good.
Russian:
Tebe stoit pojti.
you-dative cost-impersonal go-infinitive.
Lit: To you costs to go.

TMR

clause-1
  head: go-1
  agent: *hearer*
  destination: *unknown*
  aspect: phase: none
    duration: *unspecified*
    iteration: single

attitude-1
  type: deontic
  value: 0.8-1.0[1]
  scope: clause-1
  attributed-to: *speaker*
  time: time-of-speech

relation-1
  type: temporal-before
  from: time-of-speech
  to: time-of(clause-1)

	English	Japanese	Russian
Syntactic Structure	PREDICATE: had better SUBJECT: you COMPLEMENT: PREDICATE: go SUBJECT: you	PREDICATE: ii (good) SUBJ: hoo (alternative) REL-CLAUSE: PREDICATE: itta (go-past) SUBJECT: pronoun	PREDICATE: stoit' (cost)[3] SUBJECT: tebe (you-dative)[3] COMPLEMENT: PREDICATE: pojti (go) SUBJECT: tebe
Semantic Role Assignment	HAD-BETTER proposition: GO *agent: you*	II (good) theme: hoo (alternative) adjunct: itta (go) agent: pronoun	STOIT' proposition: POJTI agent: you
Lexicon: Syntactic Structure	HAD-BETTER predicate: [0] tense: past adverb: better subject: [1] complement: [2] subject: [1] inflection: infinitive	HOO[2] predicate: [1] root: (OR ii, tanosii, etc.) *tense: non-past* subject: [0] relative-clause: [2] tense: past	STOIT' predicate: [0] stoit' subject: [1] *case: dative* complement: [2] subject: [1] *inflection: infinitive*
Lexicon: Semantic Role Assignment	proposition: [2]	theme: [0]	proposition: [2]
Lexicon: Mapping to TMR	clause: [3] head: meaning-of ([2]) attitude: [4] type: deontic value: 0.8-1.0[1] scope: [3] attributed-to: *speaker* time: time-of-speech relation: [5] type: temporal-before from: time-of-speech to: time-of ([3])	clause: [3] head: meaning-of ([2]) attitude: [4] type: deontic value: 0.8-1.0[1] scope: [3] *attributed-to: *speaker** time: time-of-speech relation:[5] type: temporal-before from: time-of-speech to: time-of ([3])	clause: [3] head: meaning-of ([2]) attitude: [4] *type: deontic* value: 0.8-1.0[1] scope: [3] attributed-to: *speaker* time: time-of-speech relation: [5] *type: temporal-before* from: time-of-speech to: time-of ([3])

[1] fairly high
[2] This is an entry for a specific construction involving the noun *hoo*. There are other entries for other uses of *hoo*. The constructional illustrated here is a convention expression of modality.
[3] We are taking *tebe* to be a non-nominative subject, and the verb *stoit* to show impersonal agreement typical with non-nominative subjects. Other analyses are possible.
There are a number of paraphrases of the Russian sentence which seem to be equally conventional. For instance,
Tebe xorosho by pojti (You+Dat well counterfactual go+Inf)
Tebe nado by pojti (You+Dat necessary counterfactual go+Inf)
In fact, when we talk about conventionality we might have to operate with sets similar to the above instead of trying to find differences among each and every conventional way of realizing a certain meaning. This approach probably has its roots in the pioneering work of Apresjan (e.g., Apresjan, 1974).

- Mikrokosmos is amenable to working with incomplete information. If not all of the input conditions of the rules are present, some findings will still be possible. This property is important because we intend to deal with real texts, and we cannot hope that complete knowledge will be available. In the absence of specific knowledge, Mikrokosmos falls back on probabilistic and statistical devices.
- An important factor in the design of the microtheories is the identification of forms (above the lexical level) that are associated with some aspect of meaning *by convention*, rather than through compositional or productive rules. We follow Fillmore et al., 1988 in adopting the construction as a basic unit of analysis.

In conclusion, note how the examples in Figures 2, 3 and 4 relate to the above background assumptions of Mikrokosmos. The examples illustrate how SDLS is used as a source of clues for various microtheories, including that of lexical-semantic dependency, aspect, modality, speech acts, etc. The major finding of this paper is that TMRs are not identical to SDLS output structures, but that the latter are still necessary in that they are essential for the extraction of meaning from a text. The examples also illustrate the complex interaction of the various clues (Horiguchi 1993). For instance, the Japanese verb *morau* can signal a request-action speech act but only if it appears in a specific morpho-syntactic environment (non-past, question, speaker is subject, hearer is second object). In this environment, other clues take on special meanings. For example negation and potentiality serve only to soften the assertiveness of the request. Conventionality is also illustrated in the above examples. Many of the examples illustrate constructions that are associated with semantic and pragmatic meanings by convention. We leave the issues of non-stratificationality and working with incomplete information for future papers which deal primarily with the control structure of Mikrokosmos.

Another important contribution of this paper is to suggest a framework in which MT divergences are handled using only the mechanisms that are needed for non-divergent sentences. Our theory predicts that divergences will arise because the same element of meaning in different languages will not necessarily be expressed with isomorphic syntax, morphology, and lexical items. The Mikrokosmos TMR and the set of microtheories for all the relevant languages naturally handle the so-called divergences without any additional mechanisms.

References

[1] Bresnan, J. 1982. **The Mental Representation of Grammatical Relations** Cambridge, MA: MIT Press.

[2] Carbonell, J., T. Mitamura and E. Nyberg. 1992. The KANT MT Project. Proceedings of TMI-92. Montreal.

[3] Dorr, B. 1993. The Use of Lexical Semantics in Interlingual Machine Translation. *Machine Translation*, 7:135-194.

[4] Dorr, B. 1992. Classification of machine translation divergences and a proposed solution. *Computational Linguistics*.

[5] Goodman, K. and S. Nirenburg (eds.) 1992. **KBMT-89: A Case Study in Knowledge-Based Machine Translation.** San Mateo: Morgan Kaufmann.

[6] Fillmore, C., P. Kay and M.C. O'Connor. 1988. Regularity and Idiomaticity in Grammatical Constructions: the Case of *Let Alone*. *Language*, **64**: 501-38.

[7] Fillmore, C. and P. Kay. 1992. Construction Grammar. Course Materials. University of California at Berkeley.

[8] Hirst, G. 1987. **Semantic Interpretation and Resolution of Ambiguity**. Cambridge University Press.

[9] Hobbs, J. 1991. Interpretation as Abduction. Proceedings of ACL-91.

[10] Horiguchi, K. 1993. Extraction of Pragmatic Information in a CALL System. NLP III Project Report, Carnegie Mellon University.

[11] Jackendoff, R. 1983. **Semantics and Cognition**. Cambridge, MA: MIT Press.

[12] Levin, B. 1993. **English Verb Classes and Alternations: A Preliminary Investigation.** Chicago: The University of Chicago Press.

[13] Levin, L. and S. Nirenburg. 1994 (to appear). Construction-Based MT Lexicons. In M. Palmer, (ed.) .....

[14] Mahmoud, A.T. 1989. **A Comparative Study of Middle and Inchoative Alternations in Arabic and English.** Ph.D. Dissertation. University of Pittsburgh.

[15] Matsumoto, Y. 1992. **On the Wordhood of Japanese Complex Predicates**. Ph.D. Dissertation, Stanford University.

[16] Mel'čuk, I.A. 1981. Meaning-Text Models: A Recent Trend in Soviet Linguistics. *The Annual Review of Anthropology*.

[17] Meyer, I., B. Onyshkevych and L. Carlson. 1990. Lexicographic Principles and Design for Knowledge-Based Machine Translation. CMU CMT Technical Report 90-118.

[18] Mitamura, T. 1989. **The Hierarchical Organization of Predicate Frames for Interpretive Mapping in Natural Language Processing**. Ph.D. Dissertation. University of Pittsburgh.

[19] Nirenburg, S., J. Carbonell, M. Tomita and K. Goodman. 1992. **Knowledge-Based Machine Translation.** San Mateo, CA: Morgan Kaufmann.

[20] Nirenburg, S. and L. Levin. 1992. Syntax-Driven and Ontology-Driven Lexical Semantics. In J. Pustejovsky and S. Bergler (eds.) **Lexical Semantics and Knowledge Representation.** Berlin: Springer-Verlag. Pages 5-20.