Machine Translation: A Cognitive Linguistics Approach

Nili Mandelblit

University of California, San Diego

Department of Cognitive Science, 0515

La Jolla, CA 92037, USA.

e-mail: mandelbl@cogsci.ucsd.edu

Abstract

This paper describes a novel approach to the construction of an interlingua-based MT system. This approach emerges from the developing field of *Cognitive Linguistics.* The objective of this paper is to address one of the main problems in MT systems (besides the source text ambiguity problem): the *apparent* complexity and irregularity of translation data. The author suggests that the key to a coherent account for some of the over-diversified translation data lies in the integration of an additional knowledge-base into the translation system: the knowledge of *language usage.* This knowledge is not part of the structure of language itself but derives from cognitive mechanisms which language acts upon. Cognitive Linguistics introduces a framework which gives a central role, in any understanding of semantics, to various kinds of schemas which are triggered by the language but which are not an explicit part of language itself. The paper examines some intriguing examples of English grammatical structures and their diversified translation into Hebrew. The analysis shows that by extracting the right schema from the source text, we can generate a better representation of the text's meaning which leads to easier and more accurate generation of the target text.

1. Introduction

Some of the vital problems of current MT systems result from apparently complex and irregular translation data. Polysemies and large diversity in translation require the use of contextual constraints to guide the choice of target language words. The constraints are frequently defined *ad hoc* and the learning process results too often in an incoherent or incomprehensible set of constraints.

This paper suggests an approach for MT which may overcome some of the difficulties MT systems face today. This approach emerges from the developing field of *Cognitive Linguistics* whose framework provides a coherent analysis for some of the seemingly irregular linguistic data.

Wolfram Wills, in his article "Human and Machine Translation" (1985), describes four different but complementary "memories" needed by translators for their work:

(1) knowledge of the language system;
(2) knowledge of language usage;
(3) knowledge of the world; and
(4) knowledge of the situation.

Human translators use all four types of memory. Machine Translation systems use some (but not all) of them. Current MT systems all rely on the first type, i.e. grammatical and lexical competence, and vary with regard to their integration of the third type, i.e. world-knowledge bases. Contextual knowledge (knowledge of the situation) has come to be recognized as an important factor and is currently under investigation.

The time has come to seek new theories for the integration of the "second" type of memory - *language usage.* This knowledge is not part of the structure of language itself but derives from mechanisms which language acts upon. Cognitive Linguistics introduces a framework to deal precisely with this - the cognitive processes which manipulate linguistic expressions.

2. Theoretical Framework: the Cognitive Linguistics Approach

One of the organizing principles of cognitive linguistics is the non-autonomy of linguistic structure (Langacker, 1993). While 'traditional' linguistics studies natural language as an autonomous structure, cognitive linguistics puts the emphasis on the role played by general cognitive abilities and experientially derived cognitive models in language processing.

Traditional linguistics studies natural language as sets of strings whose meaning can be computed from the compositional structure of the language itself. Cognitive linguistics instead gives a central role, in any understanding of semantics, to various kinds of cognitive constructions which are triggered by the language but which are not an explicit part of language itself. Language according to this view plays a different role. Linguistic expressions convey minimum conventional information which triggers the appropriate cognitive constructions necessary for interpretation.

Cognitive constructions have been argued to emerge chiefly from humanly relevant experience such as our bodily movements through space, our manipulation of objects, and our perceptual interactions (cf. the *Image Schemas* of Johnson 1987, or the *conceptual archetypes* of Langacker 1991). These constructions are conceived as tools for organizing our comprehension and can structure (indefinitely) many perceptions, images, and events.

In recent years, cognitive linguists have found strong linguistic evidence for the existence of such constructs. Examples include contemporary research on the role of metaphorical mappings (Lakoff & Johnson 1980), correspondences across "mental spaces" (Fauconnier 1985), and the extraction of semantic schemata from grammatical structures (Langacker 1987, 1991, 1993; Goldberg 1992).

The point in all these studies is that the domains needed in order to understand language functioning are not just in the combinatorial structure of language itself but also in the cognitive constructions evoked by the language. These cognitive constructions are required before the content of the sentence can be interpreted, let alone translated.

3. How does Cognitive Linguistics Tie in with the Translation Process?

If language itself is only conveying minimal information, then the linguistic data of the source text might not be sufficient for full understanding of the source text and the generation of the target text. Thus, the translation process must *augment* the linguistic information with the *cognitive schemas* triggered by language. These schemas are a better representation of the meaning of the source text and can serve as a basis for target language generation.

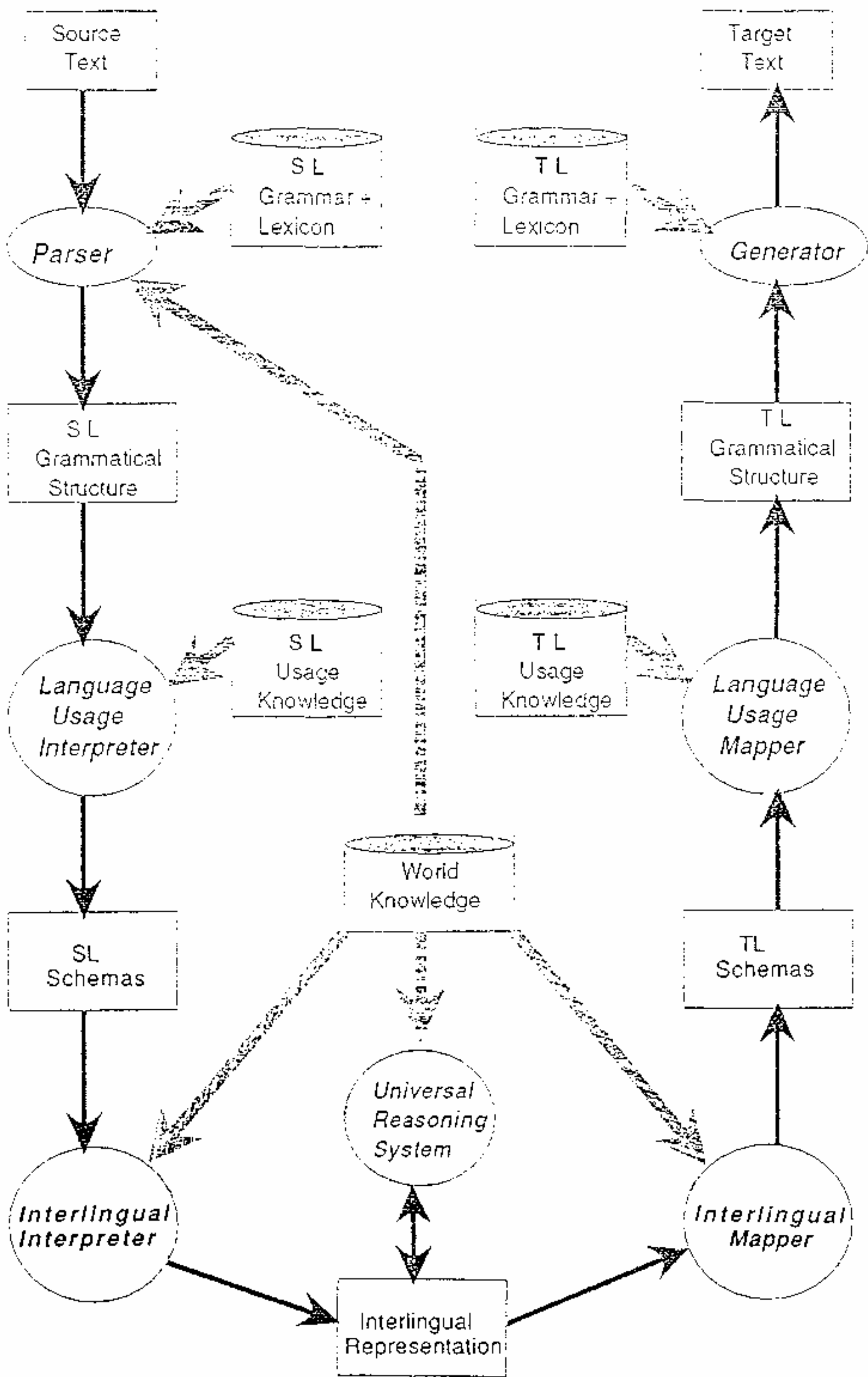
The AI approach to MT has already recognized the need for external knowledge (knowledge not included directly in the text) to bridge the gap that lies between linguistic-transcription-level MT output and a quality translation by human translators. The examples that Bar-Hillel pointed to in his criticism of MT in 1960 are precisely the problems of text understanding that research in AI addresses. Bar Hillel's argument that many semantic problems are resolved by a human translator's use of general encyclopedic world-knowledge, is dealt with in AI by integrating large databases of *language-independent domain knowledge* into the MT system (see the KBMT approach, Nirenburg 1992).

Based on cognitive linguistics findings, I would like to suggest that an additional module is needed in the process of getting a complete representation of the meaning of the text. This module takes as input a syntactic structure plus the lexical items that instantiate it (an output of the parser), and based on language-specific mapping rules, produces the general semantic schema evoked by the linguistic construction. This augmented representation of the source text serves as the basis for generation of an interlingual representation (or, in a "transfer" system, the basis for generation of the target text).

The flow of control and the role of knowledge bases in such a system are described in Figure 1. The description is based on the KBMT framework (Nirenburg, 1992). In brief, the input or source language text is processed by a set of text analysis programs. Using the knowledge recorded in the source language grammar and lexicon, plus general domain-knowledge data bases, the parser produces a syntactic representation of the text. Drawing on the output of the parser and additional language-specific knowledge bases (specifying general constructions which the source language act upon; examples will be given in section 5), the Language-Usage Interpreter produces an augmented representation of the source text. This new representation is transferred to an Interlingual Interpreter which, based on language-independent ontology and domain models (world-knowledge), produces an interlingual representation. A Reasoning System draws inferences and fills the empty slots in the interlingual representation. The result is passed on to the generator suite of programs. The Mapper induces the best schematic and syntactic form to represent the meaning in the target language and a Generator generates the target text.

While the KBMT approach makes use of three types of knowledge bases: grammars, dictionaries, and world-knowledge, the approach suggested in this paper adds an additional knowledge-base, that of *language usage* . The language-usage knowledge-base is specific to each language and serves to augment the representation of expressions in that language only.

Figure 1.


Source Text
Target Text
S L Grammar + Lexicon
T L Grammar + Lexicon
Parser
Generator
S L Grammatical Structure
T L Grammatical Structure
S L Usage Knowledge
T L Usage Knowledge
Language Usage Interpreter
Language Usage Mapper
World Knowledge
SL Schemas
TL Schemas
Universal Reasoning System
Interlingual Interpreter
Interlingual Mapper
Interlingual Representation

4. Representation of Language-Usage Schemas

A linguistic expression is associated with a semantic structure (a *Language-Usage schema)* which reflects the cognitive structure triggered by the expression (see examples in the next section). These semantic structures (schemas) are the basic data-types of the "language-usage" memory. The unit of representation is a *frame,* which has a unique name and a set of property-value pairs, called *slots* and *fillers* (as in the traditional frame-type representation in AI). The system uses the same representation language for the schemas evoked in both source and target languages, a fact which makes it possible to use reversible linguistic descriptions that can support both understanding and generation.

The set of language-usage frames and the set of universal-knowledge frames differ greatly in total number and in the elaboration of each frame. Universal knowledge frames, such as the ones used in AI (e.g. Schank 1975; Schank and Abelson 1977) are intended to capture useful chunks of encyclopedic knowledge. Thus, they are expected to be quite numerous and to be rich in content. The language-usage schemas, on the other hand, are expected to draw on a *finite* set of possible event-types (basic and recurrent human experience) and have a much more abstract content as they are intended to represent the commonality inherent in multiple experiences (Johnson 1987; Langacker 1991).

The Language-Usage schemas are encoded as frames in the same representation language as the interlingua expressions. However, while the vocabulary of the interlingua expression is that of language-independent *semantic concepts,* the vocabulary that instantiates the language-specific schemas is that of the language's own *lexicon.* The grammatical roles of the linguistic expression are mapped to particular roles (slots) in the Language-Usage schema. Some of these slots are instantiated by lexical items from the linguistic expression itself, and the fillers of other slots need to be inferred from the system's knowledge-bases (see examples in the next section).

5. Analysis of two Common English Constructions

The present section examines some intriguing examples of English grammatical structures whose translation cause difficulties in current MT systems.

5.1) Example I: The Caused-Motion Construction:

Adele Goldberg (1992) analyzes the semantics of several particular constructions in English. One of them is what Goldberg calls the *Caused-Motion construction.*

Examples of this linguistic construction include:

(1) The audience laughed the poor guy out of the room.

(2) Frank sneezed the napkin off the table.

(3) (In the last Star Trek episode), there was a woman who could think people into a different galaxy.

The form of the construction is:

[NP [V NP PP]]

where: - V is a non-stative verb

- PP is a directional phrase.

The basic sense of this construction is argued to be: 'X causes Y to move Z'.

As Goldberg shows, this syntactic form has meaning which is independent of the particular lexical items which instantiate the structure.

For example, in sentence (1), the semantics associated with this expression (i.e. 'they caused the poor guy to leave the room by (as a result of) laughing at him') is not compositionally derived from the lexical items. That is, the causal-motion sense is not expressed in any of the lexical items by itself.

The same is true for the other two examples. In each one, there is a caused-motion sense, which is not expressed by any of the individual items. The meaning of each sentence is not a combinatorial computation of the meanings of its parts. This structure is productive (see example 3 for novel usage), and its meaning is claimed to be triggered by the language speakers upon hearing the linguistic structure.

A generic Caused-Motion schema can represent the commonality in the meaning triggered by these sentences:

CAUSED MOTION Frame:

Agent:	*the audience, Frank, ...*
Object:	*the guy, the napkin, ...*
Direction:	*off the table, into another galaxy, ...*
Means:	*by) laughing, sneezing, ...*

The schema reflects an event-type in which an agent causes an object to move (or change location) in a certain direction by specific means. The grammatical roles in the linguistic expression (defined as [SUB [V OBJ PP]]) are mapped respectively to the particular roles (slots) in the frame.

When we look at the translation of the English examples (1-3) into Hebrew we see the Caused-Motion schema explicitly expressed.

Below are the English examples (E) followed by their counterparts in Hebrew (H), and a word-to-word translation of the Hebrew version into English (E*).

1) E: The audience **laughed** the poor guy off the stage.
H: Hakahal hivrich et habahur hamisken min habama betzhoko.
E*: the-audience **chased-away** (PREP) the-guy poor off the-stage **by-laughter**.

2) E: Frank **sneezed** the napkin off the table.
H: Frank heiif et hamapit min hashulchan behitatsho.
E*: Frank **threw** (PREP) the-napkin off the-table **by-sneezing**.

3) (In the last Star Trek episode, there was)
E: a woman who could **think** people into a different galaxy.
H: Isha sheyachla leha'avir anashim leolam acher bemahshava.
E*: woman who-could **transfer** people to-galaxy different **by-thought**.

In all of the examples above, the sense of the main verb in the Hebrew translation is very different from that of the main verb of the English source sentence. The Hebrew main verb expresses motion, and has a clear causative sense, as manifested by the verb pattern (Hebrew verbs all consist of some vowel pattern slotted into a skeleton of consonants; the vowel pattern defines the function of the verb in the specific instance). The vowel pattern used in all the Hebrew translations above expresses a sense of causation, a fact which supports Goldberg's premise.

The reason that the structure of the Hebrew translation is so different from its English counterpart is that there is no construction in Hebrew similar to the English *Caused-Motion* one. In order to express the meaning of the English sentence in Hebrew we must literally translate the schema triggered by the English phrase (that is, the Caused-Motion schema).

A failure to recognize the meaning that the structure imposes on the lexical items will necessitate two different definitions (and translations) for each verb. For example, the verb *laugh* would be defined as expressing both senses:

(i) laugh at

(ii) chase someone by laughter.

Similar definitions will be given to *sneeze* and *think* (examples 2-3).

Current MT systems have to either encode two translations for the verb *laugh* in the bilingual dictionary (in the case of a "Transfer" system), or specify two interlingual tokens for *laugh* (in an "Interlingua" system). These two senses must be accompanied by contextual conditions to delimit their usage. The same will have to be done with any other verb which might appear in this construction (and potentially any non-stative verb might).

By adopting Goldberg's suggestion, (viz., directly associating the semantics with the argument structure), we can avoid the need to posit ad hoc new verb senses (and *translations)* that appear only with this argument structure. Triggering the Caused-Motion schema prior to the transfer process will reduce the amount of explicit learning required by the system and provide a uniform translation process.

Note that an alternative approach to account for the data presented here is possible: resort to a lexical rule which produces motion verbs out of non-motion verbs by means of a suitable semantic operation. Now, such a lexical rule may well seem more simple, but I claim it would be less efficient and will not itself tell the whole story.

A caused-motion verb derived from the verb *laugh,* for example, will have three arguments and a sense which involves something like "X causes Y to move Z by laughter". All other derived verbs (e.g., derivations of *sneeze* or *think)* will be similar in any respect to the above sense of *laugh* except for the specification of what caused the movement (i.e. *laughing* or *sneezing).* Now, notice that the counterpart to the caused-motion verb in the Hebrew translation (i.e. the main verb in the Hebrew sentence) depends mainly on the X-Y-Z arguments of the derived verb (i.e. the moving object and the direction of movement) and not necessarily on the specific causing act *(laughing* or *sneezing).* For example, to decide whether to translate the caused motion construction using *hivrich* (chase-away) or *heiif* (throw), one has to consider the properties of the moving object (i.e. how much the moving object is active, has "free will", and can control its own movements). These considerations would be *the* same for any verb which appears in the Caused-Motion construction (whether it's *laugh, sneeze,* or *think),* and moreover, these considerations will be relevant to the verb's translation *only* when it appears in this construction. Certainly, something about the systematicity of such translation rules suggests that a more general and unified account will be not only theoretically interesting but also more efficient.

A second consideration in favor of the approach suggested in this paper is that in evoking the general schema instead of deriving an alternative sense of the verb, we save the original semantics of the verb. This way, we can account for other phenomena relating to the verb's general semantics which are encoded in the world knowledge-base.

5.2) Example II: Possessive Locutions

Langacker (1993) presents the diversity of meaning conveyed by possessive determiners in English:

(4) (a) his cup
(b) Sara's office
(c) my bus

What is the meaning conveyed by all these possessive locutions? As Langacker denotes, merely to speak of 'possession' or 'ownership' is circular and doesn't give a full account of the diversified data. Many possessive locutions (such as examples (4a)-(4c)) do not necessarily convey ownership at all nor do they convey a metaphorical extension of this basic value.

Langacker suggests that what all possessive locutions have in common is that one entity is invoked as a reference point for another. He refers to this abstract schematic characterization as the *reference-point schema (a* schema of two entities: a 'reference-point' and a 'referee'). This schema is derived from a general cognitive ability of setting asymmetrical association between two entities.

The point, for our purposes, is that interpreting locutions such as (4) is not a matter of asserting definite ownership or possession. Rather, it involves the extraction of a common frame of asymmetrical association. This asymmetrical association schema is composed, I claim, not only of the reference-point and the referee, but also of some properties which Lakoff and Johnson (1980) call "interaction properties". These are properties which emerge from interaction of objects with one another and with the world.

For example, declaring that a certain cup in a restaurant is *my cup* is probably not a declaration of possession, but an association between me and the cup which activates the fact that / *have been drinking from this cup.* The drinking activity is an "interactive property" derived from the functional properties of cups and the way people use them.

Similarly, when pointing out that a certain room is "Sara's office", what we normally mean is not necessarily that Sara owns the office, but that this office is the one *Sara is (normally) sitting in.* Again this knowledge is evoked by associating a person (Sara) with an office.

Finally, when a person says "I missed my bus", the locution *my bus* regularly denotes not a bus which this person owns but a specific scheduled bus, which the person *was planning to take.* Once more, this knowledge is derived from the regular "interaction" between persons and buses.

We should note that all the triggered senses proposed here are 'default' or 'common' ones. Different contexts may yield different interpretations (e.g., *my bus* may denote a bus I *actually own).* Since the retrieval of contextual information is, currently, beyond NLP technology, I will be dealing in this framework only with default interpretations. These nonetheless account for a large portion of source texts.

A relevant frame for the generic *reference-point* schema will be:

REF_POINT_Frame

Reference_Point:	*I, Sara...*
Referee:	*cup,* ***bus..***
Interaction:	*drinking from a cup, riding the bus...*

The universal *reference-point* schema is expressed in English using a possessive locution, but in other languages different grammatical structures might be used. If we look for the equivalent phrases in Hebrew, we will discover that examples (a) and (b) are translated to a very similar possessive structure in Hebrew. However, this is not the case with example (c).

(a)	English:	his cup
	Hebrew:	hasefel shelo
	E*:	the-cup of-he
(b)	English:	Sara's office
	Hebrew:	hamisrad shel Sara
	E*:	the-office of Sara
(c)	English:	I missed my bus.
	Hebrew:	Hifsadeti et ha'autobus shetichnanti lakachat.
	E*:	missed(I) (PREP) the-bus that-planned(I) to-take

Example (c), *my bus,* has no possessive structure equivalent in Hebrew. Hebrew just doesn't use a possessive structure to denote the association between a person and the bus that person is riding (or planning to ride). Instead, Hebrew gives a full literal description of the association, *exactly the one evoked by our cognitive frame.*

In the framework of current MT approaches, example (c) is an instance of irregularity in translation, and as such has to be specifically encoded in the transfer module. But, in fact, the translation of *my bus* into Hebrew is not exceptional. The Hebrew counterpart is simply a translation of the cognitive schema triggered by the English expression.

In an MT system capable of extracting the universal *reference-point* schema prior to the transfer phase, the Hebrew translation for *my bus* is not idiomatic or atypical. In fact, literal translation of the frame will convey just the right meaning for *any* English possessive determiner. Sometimes a shorter or more elegant target language expression may exist, but a literal translation of the frame would be sufficient to convey the meaning. The rest is a matter of *stylistic refinement.*

If we are interested in efficiency, we need of course to find out what percentage of English possessive structures is translated directly into the Hebrew possessive, and what percentage requires augmented translation. If both languages use a similar construction in most cases, it will probably be more efficient to ignore the full cognitive schema.

On the other hand, we should remember that, insofar as we are interested in multilingual systems, the extraction of the right cognitive schema is important for a multi-language generation process. The distribution of the use of the surface possessive structure in expressing the reference-point relationship might change from one language to another. Retrieval of the general *reference-point* schema extracts the full meaning of the locution, and serves as a more universal language-independent representation.

The main problem in deriving the *reference-point* frame is to determine the relevant "interaction properties" associated with each instance. These properties need to be derived from general *world-knowledge.* For the system to be able to assign the default Interaction property (i.e. the common interaction between the Reference-Point and the Referee), it must have a knowledge-base which contains the semantic configuration of concepts (or words). These semantic configurations should reflect not only thematic role descriptions (Fillmore 1986), but also nonlinguistic images, such as an object's constituents, shape, creation, purpose and function (for discussion, see Pustejovski 1991).

The semantic structure of the concept *cup,* for example, should specify what the artifact is used for; i.e. drinking. Wierzbicka (1984) suggests that an adequate definition of a concept must include the "prototypical situation" in which the object is meant to be used. That is, although there are many ways for people to act on a *cup* (many types of interactions), it is argued that certain relations are "privileged" in the semantics of the concept.

A partial semantic description for the concept CUP(X) would include:

CUP (X)

Function: *hold*(X*,* liquids);
drink (people, from X);

Similarly, a definition for *Office*(X) will include the fact that *Work(People, in X),* and for *Bus(X)* that *Ride(People, X).* To find the right "interaction" property, our system should consult the knowledge-base to find a predicate whose arguments are the 'reference-point' and the 'referee' of the schema.

The output of the Language-Usage Interpreter provides a frame with an empty slot (the "Interaction" slot). In a "transfer" MT system, the derivation of the right "interaction" between the reference-point and the referee will be done in the *generation* suite of programs, after mapping the 'reference-point' and the 'referee' slots to their counterparts in the target language. In an "interlingua" system (such as the one described in this paper), the source-language schema is mapped onto an interlingual representation with the 'reference-point' and the 'referee' slots mapped onto language-independent semantic concepts. The "interaction" properties, in the latter case, are derived from the universal world knowledge data-base.

6. Summary

In this paper I have described a novel approach to the construction of an interlingua-based MT system. This approach is based on the observation that many of the problems that arise in MT systems (besides the source text ambiguity problem that has been thoroughly discussed elsewhere) are a result of over-diversified translation data. Cognitive Linguistics suggests a coherent account for some of these apparently irregular data by taking into account the cognitive schemas evoked by the linguistic expressions during language processing. Adding a module to the MT system which triggers the right language-specific schemas prior to target-language generation can provide a more accurate and uniform translation process.

References:

Bar-Hillel, Y. 1960.
The Present Status of Automatic Translation of Languages. In F.L. Alt (ed.), *Advances in Computers,* Volume I. Academic Press.

Fauconnier, G. 1985.
Mental Spaces. Cambridge, Mass: MIT Press.

Fillmore, C. 1968.
The Case for Case. In Bach, E. & Harms, R.T. (ed). *Universals of Linguistic Theory.* New-York, NY: Holt, Rinehart and Winston:1-88.

Goldberg, A. 1992.
Argument Structure Constructions. Berkeley: University of California. Doctoral Dissertation.

Johnson, M. 1987.
The Body in the Mind: The Bodily Basis of Meaning, Imagination, and Reason. Chicago and London: University of Chicago Press.

Lakoff, G. & Johnson, M. 1980.
Metaphors we live by. Chicago: University of Chicago Press.

Langacker, R. 1987.
Foundations of Cognitive Grammar. Volume I. Stanford University Press.

Langacker, R. 1991.
Foundations of Cognitive Grammar. Volume II. Stanford University Press.

Langacker, R. 1993.
Reference Point Constructions. *Cognitive Linguistics.* 4.1-38.

Nirenburg, S., Carbonell, J., Tomita, M., & Goodman, K. 1992.
Machine Translation: A Knowledge-Based Approach. Morgan Kaufmann Publishers.

Pustejovky, J. 1991.
The Generative Lexicon. *Computational Linguistics,* 17(4):409-441.

Schank, R. 1975.
Conceptual Information Processing . Amsterdam, Holland

Schank, R. & Abelson, R. 1977.
Scripts, Plans, Goals and Understanding. Hillsdale, N.J.

Wierzbicka, A. 1984.
Cups and Mugs: Lexicography and Conceptual Analysis. *Australian Journal of Linguistics* 4.025-255.

Wilss, W. 1985.
Human and Machine Translation. In *Translation Perspectives II,* National Resource center for Translation and Interpretation, SUNY-Binghamton.