Treatment of Meaning in MT Systems

Sergei Nirenburg and Kenneth Goodman

Center for Machine Translation, Carnegie Mellon University

ABSTRACT. Fully-automated, high-quality MT systems can be developed only if a realistic way to treat meaning in natural languages is found. This paper analyzes a set of positions on the treatment of meaning in machine translation, which we consider the crucial point of difference among MT approaches. We present and criticize a chain of claims that is skeptical with respect to the prospects for meaning-based machine translation. Viewed as a meta-stance, this chain of opinions argues — wrongly, we suggest — against the feasibility, the need for and even the possibility of meaning analysis as the basis for machine translation technology. We set out to demonstrate that these viewpoints, however popular, are flawed. In fact, we suggest that the methodological debate between the interlingua and the transfer approaches to MT is, in fact, misplaced. The meaning-based approach to MT can in principle be explored using either of these architectures or — even more promisingly — some combination of them.

1. Introduction

At the methodological level, polemics about the state of the MT art are most often couched in terms of the differences and tensions between the transfer and the interlingua approaches. Over the years, a great deal of folklore has accumulated about the pros and cons of each of these MT paradigms. We believe that the discourse on this topic is not as organized as it should be. A number of claims and opinions on the subject have been made publicly. Unfortunately, most of them appeared only in prefaces and introductions to books and articles on MT. After mentioning the predominant methodological issue on page 1 and briefly identifying their own positions, MT authors typically plunge into descriptions of their own systems or models without further analysis of the methodological issues. Under these circumstances, it is not surprising that the methodological argument is not conducted at an adequate level of detail.

Even in discussions devoted specifically to the "transfer vs interlingua" issue (such as, for instance, the panel on this topic at the UMIST 1989 MT workshop) many of the arguments remain too general and iconic. As a result, a discrepancy can be detected between the methodological beliefs held by MT practitioners and the actual (theoretical and practical) preferences and results in the field. Even at the methodological level, criticism is often directed at opinions that are not, in fact, held or defended by one's opponents.

Judgments about paradigms may differ depending on the specific profile of a system. MT is simultaneously an empirical discipline and a technological pursuit. Depending on the primary direction of research and development in a project, different criteria should be used to evaluate the utility and quality of systems developed in it. There are i) production systems, ii) production system prototypes, iii) proof-of-concept systems which demonstrate the utility of a theoretical or descriptive approach to MT or a component process in MT (e.g., syntactic analysis, treatment of referential meaning, etc.) and iv) technological testbeds for producing MT systems (including specialized knowledge acquisition interfaces, debugging tools, control environments, etc.).

One must also distinguish between evaluations of particular projects and evaluations of entire approaches. If it is claimed that Project A used Approach X and failed, it does not necessarily follow that Approach X is bad. Similarly, the claim that Project B used Approach Y and succeeded does not in itself mean that Approach Y is superior. One reason for this caution is that large MT projects tend to feature elements from several MT paradigms. Therefore, it is often a gross generalization to call a particular project purely interlingual or purely transfer. A finer-grain taxonomy of MT approaches is needed. This is a central methodological point of this paper.

Actually, the components of an interlingua text are produced by and informed by several interconnecting subsystems. In our knowledge-based MT system, KBMT-89 (Goodman and Nirenburg (1989)), the interlingua is created by an analyzer that consists of a set of programs and knowledge sources, including source-language lexicons and grammars, mapping rules for syntactic features and structures, and an ontology or domain model. The generation side of the triangle is equally complex.

We would like to argue that the ultimate point of contention in methodological debates among the MT researchers is not so much the differences between the transfer and interlingua approaches but rather the attitude to treatment of meaning. It so happens that, as a rule, those MT workers who de-emphasize the importance of meaning extraction tend to favor transfer-oriented systems, while those who insist on understanding as a prerequisite of translation tend to prefer interlingua-oriented ones. The reasoning leading' to this latter preference can be clarified using an example of a research program in meaning-oriented MT, namely KBMT-89.

In very general terms, our research-and-development activity can be characterized as follows. Its methodological basis is meaning-oriented MT in an interlingua paradigm.[1] The Center for Machine Translation's KBMT systems are research-oriented systems that come under the rubrics "proof of concept" and "technological testbed." A great deal of attention is paid to the functionality range of software engineering, including architecture and control, as well as to the massive task of knowledge acquisition. A new version of the system is developed, tested and demonstrated, on average, annually. The systems we build and demonstrate are gradual approximations of an ideal interlingua MT system. At present, although some facets of our systems are relatively complete and stable (some of the grammars, parsers, the integrating control structure), we have only partially accounted for many others, such as, for instance, many areas of domain knowledge, the lexis of the languages involved and some of the heuristic rules ("microtheories") used for treating particular linguistic phenomena. Systems developed at CMT differ mainly in the amount of knowledge that has been accumulated for use in them. They share a number of important characteristics.

The flow of control in these systems is as follows. The input text is processed by a battery of text analysis programs. Using the knowledge recorded in the SL grammar and lexicon, these programs (after several stages of processing) produce an expression (usually called *interlingua text* or ILT) in a specially defined textual meaning representation language. Elements of ILT are produced based on lexical, grammatical and pragmatic meanings extracted from the source text. Some ILT elements are instances of concepts in a domain model. Some others are values of various semantic and pragmatic properties suggested as necessary

[1] This statement will immediately instantiate the familiar triangular diagram in the mind of the reader. It is worth remarking that this quasi-standard illustration of the interlingua system design (and the equally ubiquitous trapezoid or rectangular diagram of the transfer system design) can be confusing or misleading. The use of circles, arrows and labels to represent the top-level structure of a system should properly be understood as the creation of a son of visual slogan. There is much detail that does not easily lend itself to graphical representation. In simple sketches of interlingua systems, for instance, the source language is seen at the left and giving rise via an arrow to an "interlingua" that in turn points with an arrow to the target language. In simple sketches of the transfer architecture (cf. the figure below) it is quite difficult to express the actual nature of the transfer structures.

components of ILT. Very generally.[2] ILTs are hierarchical structures of clause-level representation units connected through domain and textual relations from a predefined set and characterized by speaker attitude values.

At present, we cannot produce a complete ILT fully automatically. Therefore, our systems use an interactive "augmentor" (the concept of such a program was first demonstrated by Kay (1973)). As our knowledge about language processing grows, we expect the role of the augmentor to diminish.

A complete ILT produced jointly by the analyzer and the augmentor is passed on to the generator suite of programs, which includes a text planner, a lexical selection module and a syntactic realizer. The generator uses a TL lexicon and grammar and other heuristic knowledge sources necessary for generation.

The number and nature of microtheories used for extracting and representing meaning components varies among system instances in our project. Microtheories are not necessarily "homegrown." They can be routinely imported and adapted. Thus, some of the heuristics can, in principle, depend on some surface phenomena in the source text. If such heuristics are used, they constitute a "transfer" element in a generally interlingual system.[3] We have no objection in principle to a transfer-oriented system as long as it incorporates a meaning treatment module. It seems, though, that at least for purely technological reasons it is simpler to formulate polysemy resolution rules in terms of domain model elements rather than lexical units of the source and target language, as is the practice in transfer systems that deal with polysemy resolution at all (e.g., SPANAM, Vasconcellos and Leon (1985)).

Meaning-based MT offers additional scientific incentives. It is a paradigmatic task for computational linguistics in that components necessary for a knowledge-based MT system are also necessary components in practically every other application system dealing with automatic processing of natural language. MT is, therefore, one of the most attractive comprehensive applications for various computational-linguistic theories — morphological, syntactic, semantic and pragmatic. Historically, "pure" transfer systems have not addressed the problem of semantic and pragmatic ambiguity resolution as a central problem for machine translation. These approaches are based on recognizing meaning without representing it (other than in terms of target language lexical units). Methodologically, this is a result of i) expecting very small amounts of ambiguity and ii) relying on similarities between the source and target languages to try to "preserve" in translation any ambiguities that appear in the source text. When (or, in fact, if) a transfer system has any theoretical connections, they are predominantly connections to theories of syntax. Of course, such theories constitute the bulk of theoretical work in modem computational linguistics. But this state of affairs does not make analysis of meaning less essential. On the contrary, it becomes even more important for semantic theorists to have an adequate testbed for their ideas. And meaning-oriented MT provides just that.

The purpose of the above discussion of our view of meaning-based MT is not to give a complete or adequate account of its workings. As such, it certainly fails. Our intent is to illustrate the complexity of system organization, which cannot be readily summarized in the triangular icon of the interlingua approach. In a narrow sense, the "interlingua" in our systems is the set of all well-formed expressions (ILTs) in the

[2] A detailed description of our text meaning representation language, TAMERLAN, is given in Nirenburg and Defrise, forthcoming.

[3] One such heuristic is the decision to retain in the target text the boundaries of sentences in the source text. It is well known to human translators that translations can be improved if one can combine some of the source sentences together or break some of them into several sentences. If knowledge is available for judging when such actions are appropriate, a meaning-based MT system can use it to determine sentence boundaries in the target text. But if this microtheory is not yet available, a good working heuristic is to copy the boundaries in the source. Other phenomena are treated in meaning-based systems in ways similar to their treatment in transfer systems. One such phenomenon is the handling of unambiguous terminological lexis (such as, for instance, chemical nomenclature).

textual meaning representation language. In a broader sense, it should also include the domain model underlying (most entries in) the lexicons. It is in terms of the expressive power of this interlingua (and the feasibility of the "microtheories") that the quality of our version of the meaning-based approach to MT should be discussed. Other views of the treatment of meaning in MT should be judged on the quality of their own representations and analyses.

In what follows, we discuss several opinions that are frequently put forward in arguments about meaning-oriented MT. The following list summarizes these opinions and puts them in a logical chain of arguments which goes from extremely strong and general criticisms toward more specific and limited ones. The list is by no means complete. We hope only that it is representative. After presenting the list of criticisms, we will evaluate each in turn and in greater or lesser detail.[4]

1. Translation is *not possible;* if it is, then
2. Meaning is *not required* for translation; if it is, then
3. Meaning is *not definable;* if it is, then
4. Meaning in different languages is different and *not compatible;* if it is compatible, then
5. One *cannot represent* this meaning in a language-independent way:
    - the language of representation will be heavily slanted toward one particular natural language;
    - it is difficult to come up with the necessary set of language-independent primitives and to ensure completeness of meaning representation.

   Furthermore,
6. It is not possible to base meaning representations on a complete logical calculus. Therefore, one can never *prove* the correctness of any representation, particularly that it is free of contradiction; or that the same meanings will be always represented similarly. If constraints of this sort are demonstrated to be manageable or unnecessary, then
7. It is impossible to ensure that the meaning can *actually be extracted* from the source-language text and rendered in the representation language; at least, it is not possible to extract meaning *completely automatically.*
8. Even if meaning-based translation systems can be built, they will produce not translations but rather paraphrases of source language texts.

2. Possibility of Translation

A long and rich history attaches to philosophical arguments against the possibility of translation as a re-representation of meaning. Of course, the *locus classicus* is Quine's *Word and Object* and his theory of radical translation. The point of radical translation was that "manuals for translating one language into another can be set up in divergent ways, all compatible with the totality of speech dispositions, yet incompatible with one another. In countless places they will diverge in giving, as their respective translations of a sentence of the

[4] Discussions of points 2 and 3 are combined in Section 3 below. The other criticisms are considered in separate sections.

one language, sentences of the other language which stand to each other in no plausible sort of equivalence however loose..." (Quine, 1960:27).

Quine's behavioristic "translation manuals" may be understood loosely as analogues of the grammars, lexicons and programs of machine translation. The idea of radical translation was not that it is impossible to translate natural languages — humans do it all the time — but that what is translated is not the "same meaning." In other words, there are no meanings *qua* meanings to translate: The alleged absence of independent identity conditions for meanings entails that there are no language-neutral semantic entities. We examine this and related philosophical issues in Goodman and Nirenburg (1990), and will not develop the points here.[5] Suffice it to say that, following Katz (1988), we believe (i) Quine was just mistaken and (ii) it is not in any event clear how to apply his arguments to machine translation. The goal of Quine and some of his allies is mainly to demonstrate the underdetermination of scientific theories by evidence and it would be specious to transport the issues and arguments too quickly to our domain.

To contend that translation is not possible, then, is on one reading just false — and should be uncontroversially so. On another reading it entails, in some cases informally, a group of critiques of the "meaning" relation; and so it is to those critiques that we turn next.

3. Understanding and Translation

It has been suggested that meaning is not required for machine translation. The idea is that a source-language sentence might be translated automatically into a target-language sentence by *statistical* means. The idea is as old as MT itself and attracted Warren Weaver in the 1940s and informed the early approaches at RAND and the National Bureau of Standards through the early 1960s (see de Roeck (1987) and references cited therein).

Most recently, Brown *et al.* (1988) report on experiments with a statistical approach to machine translation which "... eschews the use of an intermediate mechanism (language) that would encode the 'meaning' of the source text." The contention in this approach is that "... translation ought to be based on a complex glossary of correspondence of fixed locutions" and, more fully,

> 'Translation can be somewhat naively regarded as a three stage process:
>
> (1) Partition the source text into a set of fixed locutions.
>
> (2) Use the glossary plus contextual information to select the corresponding set of fixed locutions in the target language.
>
> (3) Arrange the words of the target fixed locutions into a sequence that forms the target sentence."

In other words, language in this approach is treated not as a productive system but as a fixed and unproductive set of canned locutions.

The applicability of an MT system built according to this approach is restricted to the cases where there are vast textual corpora of translation equivalents. But even when such materials are available, completely uninterpreted comparison will lead to errors simply because the human translators who produced

[5] Note that our meta-argumentative chain appears in an abbreviated form in the companion paper to plot out philosophical issues related to the differences between the transfer and interlingua MT approaches.

the translations in the corpus in the first place do not translate word-for-word or even sentence-for-sentence. The meaning expressed by a lexical unit in the source language can be rendered as an affix or as a syntactic construction in the target language. Nagao (1989:6-7) writes: "... although they are infrequently used in European languages, in Japanese there are many words of respect and politeness which reflect the social positions of the speakers, as well as distinctly male or female expressions which lie at the heart of Japanese culture. These are factors which must be considered when translating between Japanese and European languages ... Even if those factors are not explicitly expressed in the target language, they should be inferable from the context, from the psychological state of the speaker, or from the cultural background of the language." It will be difficult for a purely statistical system to detect such phenomena.

A major shortcoming of the statistical approach is as follows. What, that is, does one do when in a certain text the English word *lead* is translated into Russian as *provod* ("cable") 17 times and as *svinets* (the metal) six times? Can we, indeed, be democratic and go with the greater number of votes? Clearly not. Therefore, according to the statistical approach, one has to make step 2 in the above definition of the translation process a conditional one. The conditions will have to be formulated in terms of the "contextual information," that is, in terms of lexical units as such, syntactic structures, or lexical and other meanings. Depending on the particular choice from this list, the statistical approach will, we suggest, rediscover direct, transfer or interlingua models of MT.

While it does not seem that a purely statistical approach is adequate to the task of MT, we believe that a statistical component can be very useful in a practical MT system, both as an aid in knowledge acquisition and as a way of testing meaning preferences.

That meaning understanding is not necessary is also maintained by another group of researchers who observe that, for instance, the polysemous Spanish noun *centro* is translated into German as *zentrum* no matter which of the senses of *centro* was used in the SL text (see below). The question then is, Why waste time detecting and representing the meaning of the input string when the target language correlate is always the same? Similar claims have been made about syntactic ambiguities (e.g., Pericliev (1984)) and ambiguities of prepositional phrase attachment (e.g., Kay (1989)).

A typical formulation of this position is given by Ben Ari et al. (1988:2): "It must be kept in mind that the translation process does not necessarily require full understanding of the text. Many ambiguities may be preserved during translation [Pericliev 84], and thus should not be presented to the user (human translator) for resolution."

Similarly, Isabelle and Bourbeau (1985:21) contend that, "Sometimes, it is possible to ignore certain ambiguities, in the hope that the same ambiguities will carry over in translation. This is particularly true in systems like TAUM-AVIATION that deal with only one pair of closely related languages. The difficult problem of prepositional phrase attachment, for example, is frequently bypassed in this way. Generally speaking, however, analysis is aimed at producing an unambiguous intermediate representation."

This position is, in fact, a system-completeness argument. What it says is that, for a given SL - TL pair and (i) a given set of dictionary senses of each SL word and (ii) recognized SL syntactic patterns, there will be cases in which all the senses of a SL lexical unit will be realized by a single lexical unit in the TL, or an SL syntactic construction can be re-created without change in TL. The familiar sentence

(1) I saw a man on the hill with a telescope

can be translated into some languages without the need to understand the dependency characteristics of the

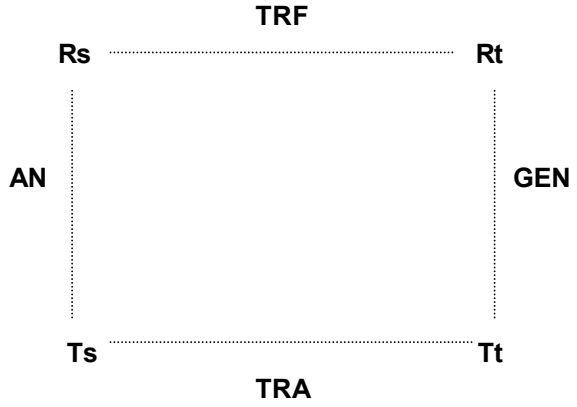
prepositional phrases, just by stringing them, in their original order, after the direct object. This type of knowledge allows the system builders to keep the sizes of lexicons and grammars smaller.

The set of arguments about preserving ambiguity has serious limitations. It is clear, for instance, that *centro ciudad* should be translated in colloquial English not as "town center" but as "downtown." Also, it is only possible correctly to render (1) in, say, Russian, if one understood the prepositional attachments. Otherwise it will not be possible to select prepositions and casual forms adequately. The argument in Pericliev (1984) is supported by a manual analysis of 200 short English phrases and their translations into Bulgarian. No syntactic ambiguity was found in about 150 of these phrases. In about 25 cases ambiguity could be preserved by simple substitution. And the remaining 25 could not be treated this way. Therefore, Pericliev's claims lack generality.

Considering the amount of work required to put together a non-trivial MT system, it is quite reasonable to strive to constrain the size of the knowledge acquisition task. At the same time, one must remember that a system strongly relying on the "ambiguity preservation" method is extremely vulnerable in situations where (i) the lexicon is growing while the system is in use or (ii) when additional languages must be introduced. Every new word sense added to the lexicon carries the potential of ruining the possibility of retaining ambiguity in translation. And this means that extra attention must be paid to the maintenance of the lexicons.

The problem of working with increasingly large dictionaries and grammars remains to be solved for all MT systems, irrespective of the theoretical approaches they follow. There are also cases (especially when very concrete technological terminology is concerned) when knowledge about the field of translation and the authoring style of a particular type of text will lead to the possibility of rendering certain elements of SL through unconditional (and, possibly, multilingual) substitution by TL counterparts. However, actual MT systems will be judged by their "maximum" capabilities in treating complex, not simple, problems.

Some of the MT literature is devoted mostly to design and metalevel (not to say MT-theoretical) issues (King (1981), Arnold and des Tombe (1987), Warwick (1987)). Typically, such contributions suggest or discuss an abstract theory of translation as a series of transformations among representations. A typical series of definitions follows:

> The following diagram forms a good basis for the study of representations in a transfer-based translation system:
>
> 
>
>
> In [the above figure], Ts and Tt are texts, where a language is regarded as a set of texts. TRA is a binary relation, consisting of pairs of texts [Ts, Tt] where Tt is a translation of Ts. So, given two languages, SL and TL, TRA ≤ SL x TL. We introduce, furthermore, ρ which is a set of representations of some kind. Rs and Rt are both members of this set. We will write R when it

is unimportant whether we are dealing with Rs or Rt, or when the context makes it clear which is intended (Johnson *et al.* (1985)).

In other abstract definitions of the translation process the number of transformations is larger, but in none of them is the question of semantic ambiguity dealt with centrally. Thus, for instance, having sketched an abstract view of the translation relation, Whitelock (1989:6) characteristically adds: "One question I have not touched on here is the question or ambiguity. What I have been talking about is a many-many relation 'possible translation' which may be computed monotonically from the axioms of the grammar and lexicon of the languages concerned. Optimally, this is viewed as a totally different question from determining the best translation, given an unbounded amount of real world knowledge, discourse context, etc. Computing this relation requires inference which is presumably defeasible."

Intrinsic in this statement is the opinion that what the understanding of meaning adds to the quality of translation is the possibility of getting to the "best" translation, as opposed, presumably, to "adequate" though not "best" one. We believe that the translation relation suggested by Whitelock has no way of guaranteeing even an "adequate" let alone the "best" translation. It has been amply and repeatedly demonstrated through multiple examples in the MT and natural language processing literature, starting at least with Bar Hillel (1960), that no adequate translation of realistic-size texts can be obtained if semantic issues are not addressed. A great deal of ingenuity and *ad hoc* knowledge acquisition is needed to avoid ambiguity resolution in MT. And, in fact, translation using this approach can be successfully achieved only for carefully selected subsets of texts, served by dictionaries in which single-sense entries predominate. This is one reason why postediting is such a necessary stage in typical direct-approach and transfer-oriented MT environments. It seems that even when discussing general design issues in MT, it does not make sense to exclude considerations of ambiguity resolution.

4. Meaning Across Languages

Some MT researchers adopt the position that different languages employ different concepts, or employ concepts differently, and this short-circuits attempts at meaning extraction. Thus Amano (1989:2) writes that "Natural languages have their own articulation of concepts according to their culture." To illustrate this point, Amano reports that where the English word *moustache* is customarily defined in English dictionaries as comprising hair on the upper lip, the Japanese *kuchi-hige* is defined in one (unspecified) Japanese dictionary as a "beard under the nose." (Actually, the ideographs for *kuchi-hige* stand for 'lip' or 'mouth' and 'whiskers'.) From this we are urged to infer that what Japanese speakers *mean* by *kuchi-hige* is somehow different than what English speakers mean by *moustache.* Of course, this opinion is simply a particularly hirsute version of Sapir-Whorfism that depends crucially on the vagaries of dictionary entries. Amano states that "natural languages have their own articulation of concepts according to their culture. Interlingua must naturally take account of this" (ibid.). But this is a misunderstanding of the concept of interlingua. What differs among languages is not the meaning representation but rather the means of realizing this meaning. The meaning of *kuchi-hige* and *moustache* will be represented in the same way in an interlingua text. The realizations of this meaning in the two languages will be different. It is in the interlingua-TL dictionary that a connection is established between an interlingual meaning representation and the language-particular linguistic expression.

This is not the place to argue against linguistic and cognitive relativism. The idea of linguistic relativity is, in fact, neutral with respect to the tasks of computational linguistics. It should be sufficient to point out that however convenient dictionaries might be as explicators of meaning for humans, it is a mistake to

appeal to them as formal indexes of a culture's conceptual structure. That is to say, even within a language many terms may be rendered in different and apparently incompatible ways. To contend that meaning exists intralingually but not interlingually is to fall prey to such examples and to slip into the meanest sort of relativism, even unto idiolects. In practice, of course, indigenous *realia* can be described encyclopedically and then assigned a linguistic sign (possibly, a direct calque from the original language).

5. Feasibility of General Meaning Representation

One argument against language-independent meaning representation, usually referred to as interlingua, is known as "cultural-imperialist." To wit, the way the interlingua is built reflects the worldview behind one dominant language. Examples of phenomena with respect to which "cultural imperialism" can be established include the cross-linguistic difference in subcategorization behavior of verbs, the grain size of concept description and the difference in attitude similar to the above *moustache* case. For instance, a single interlingua concept can be suggested to represent the main sense of the English *put* (as in *Put a book/glass on the table.* This might be considered a case of English cultural imperialism because in Russian this meaning can be expressed either as *položit' or postavit'* depending on some properties of the object *of put.*[6] Additional examples abound in the MT literature.

The granularity of a large-scale meaning representation is always influenced by linguistic data, since the acquisition of knowledge necessary to support such a representation is done by humans who are, naturally, influenced by the languages they speak and the textual corpora and human-oriented dictionaries they use to determine meaning unit boundaries. It seems that the "cultural imperialism" argument is directed at the wrong target.

The simple view of the interlingua as a representation capturing all meanings in all languages is certainly limited because it talks about an ideal approachable only asymptotically. Compare, for instance, the following statement by Nagao (1989:6): "... when the pivot language method is used, the results of the analytic stage must be in a form which can be utilized by all of the different languages into which translation is to take place ... This level of subtlety is a practical impossibility." On a more technological level, Schneider (1989:128) justifies the choice of paradigm in the METAL project as follows: "METAL employs a modified transfer approach rather than an interlingua. If a meta-language were to be used for translation purposes it would need to incorporate all possible features of many languages. That would not only be an endless task but probably a fruitless one as well. Such a system would soon become unmanageable and perhaps collapse under its own weight."

This "maximalist" view of interlingua is so popular probably because it is conceptually the simplest. In operational terms, however, it is as useful to talk about such a conception of the interlingua as about a set of bilingual dictionaries among all the language pairs in the world. A practical interlingua should be viewed both as an object and as a process. Viewed as an object, developed in a concrete project, an interlingua should be judged by the quality of the translations that it supports between all the languages for which the corresponding SL-interlingua and interlingua-TL dictionaries have been built. As a process, its success should be judged in terms of the ease with which new concepts can be added to it and existing concepts modified in view of new textual evidence (either from new languages or from those already treated in the system). In practice, all interlingual systems start with the description of the semantic (sub)realms of a small set of languages and expand only when it becomes feasible from the standpoint of project resources. This

[6] The difference can be glossed as that between *put flat* and *put upright.* A book can be "put" either way; a glass will be usually "put upright"

is true about such different interlingual systems as ATLAS-II (Uchida (1989)), Rosetta (Landsbergen (1989)) or KBMT-89 (Goodman and Nirenburg, 1989).

It is characteristic, though, that even interlingua-oriented workers find it necessary to offer qualifying explanations of their paradigmatic choices. Thus, Landsbergen (1989:85) writes:

> 1. From the point of view of the system's architecture Rosetta is clearly an interlingual system. It consists of an analysis component that translates from the source language into an intermediate language, of which the expressions are semantic representations, and a generation component that translates from this intermediate language into the target language.
> 2. On the other hand, the intermediate language of Rosetta is not a universal interlingua, but is defined for a specific set of languages. So Rosetta is not interlingual in this strict sense.
> 3. In an ideal interlingual system the analysis and generation component for each language can be developed independent of the other languages. We will not discuss here to what extent this is desirable or possible, but it is clearly not the case in Rosetta.

A system (such as Rosetta) based on the principles of meaning analysis and absence of direct correspondences between the elements of SL and TL must have the right to be called "interlingual" unapologetically. It does not seem appropriate (in fact, it looks like a double standard) to require completeness as proof of feasibility for interlinguae, while allowing adequate behavior in a limited domain for a limited set of language pairs (usually, a single language pair) to be the criterion of success of a transfer system.

Often, the argument about infeasibility of interlingual MT is presented as a "basic assumption" and not argued for, as done, for instance, by Arnold and des Tombe (1987:117): " ...the translation relation is fundamentally and irreducibly a relation between *linguistic* objects. The representation languages must be linguistic in nature, and cannot therefore be completely neutral with respect to different natural languages, in the way that a genuine interlingua would be." It is not clear what is meant by "linguistic." Is the formalism Arnold and des Tombe suggest (or formalisms based on similar principles, presented, e.g., in Johnson *et al.* (1985) or Arnold *et al.* (1988)) in any sense a more "linguistic" notation than an artificial language designed to capture textual meaning? If "linguistic" is equated with "stemming from a syntactic theory" then we strongly disagree, because, in our understanding, translation is based on mapping meanings, not syntactic structures.

The "maximalist" view of the interlingua sometimes constitutes the main reason for not selecting this approach for a particular project. This is sometimes the case even in the presence of task specifications (such as multilinguality) which suggest an interlingua approach. Thus, the reasons for not selecting this approach for the original EUROTRA project are given by King (1981) as follows: "EUROTRA tries, at its deepest level of representation, to characterize the semantic relations between constituents in the text via a set of relations based on an expanded form of case grammar ... However, since the set of relations are defined as those useful for translation and are only 'universal' within the project, there is no attempt to reach a [sic] ideal, genuinely universal semantic representation."

In reality, the EUROTRA approach evolved in such a way that many of the elements usually associated with the interlingua approach (first and foremost, analysis of meaning) are present in it (cf. Durand et al., *forthcoming).* Therefore, the traditional arguments against interlingua and for transfer approaches should perhaps be presented today as arguments against the use of meaning in translation. If meaning is not considered essential for translation, then a version of a transfer approach should be the choice, since interlingual approaches crucially depend on meaning representation. We still believe that it would be more convenient and general to couch the meaning analysis in language-independent form rather than analyze the meaning of a source language in terms of lexical units of a target language (and this is what bilingual

transfer dictionaries, in fact, do). But this argument is of a secondary nature. The main point established by this convergence in approaches is that treatment of meaning is central to the task of MT. We are, in reality, disagreeing only on ways and means.

We want at this point to discuss several well-known opinions about and criticisms of interlingual MT. It seems that most of them, indeed, refer to treatment of meaning rather than to interlingua as such.

The most well-known and large-scale early experiment with interlingual representations ended in self-admitted failure: " ... we have tried an approximation of the interlingua ('pivot') approach and found it wanting. In the... CETA system, the pivot representation was of a hybrid sort, using as vocabulary the lexical units of a given natural language, and as relations the so-called 'universals' corresponding to our current logical and semantic relations, plus abstract features such as semantic markers, abstract time and aspect and so on" (Vauquois and Boitet, 1985:35).

Design characteristics of the CETA interlingua were, in fact, drastically different from those usually associated with interlingual systems. Hutchins (1986:190f) summarizes the characteristics of the CETA pivot language as follows: "The formalism was designed primarily as an interlingua for syntactic features, i.e., as the common 'deep syntactic' base of the languages in the system ... Its lexicon, however, did not represent a common base; instead the pivot language conjoined the lexical units of whichever two languages were being processed ... In other words, while the CETA pivot language was a true interlingua in syntax, it was a bilingual 'transfer' mechanism in lexicon. Further, it was not intended that all sentences with the same meaning would be analyzed as ... one unique pivot language representation. Nevertheless, although there were thus as many 'pivot languages' as there were SL-TL pairs analysed, all shared the same syntax and in this respect CETA considered their formalism as a first step in the direction of a 'universal language.'" As described above, the design of CETA is interlingual only in name. In fact, it is practically identical to that of a standard modem transfer-based MT system!

Still, the fact that this system recognized itself as interlingual and self-admittedly failed has been used to justify objections to interlingual MT, for instance, in the METAL project: "It is frequently argued that translation should be a process of analyzing the source language into a 'deep representation' of some sort, then directly synthesizing the target language... We and others (King, 1981) contest this claim ... One objection is based on large-scale, long-term trials of the 'deep representation' approach by the CETA group at Grenoble ... After an enormous investment in time and energy, including experiments with massive amounts (400,000 words) of text, it was decided that the development of a suitable pivot language (for use in Russian-French translation) was not yet possible" (Bennett and Slocum (1985: 112)). Comparing this opinion to the above discussion of the CETA project, one has to conclude that the self-admitted failure of CETA should have raised doubts about the feasibility of the *transfer* approach rather than the interlingua one.

It is sometimes claimed that meaning representation that does not use elements of natural language is difficult to design: "It is very difficult to design [a meaning representation] in the first place, and evermore so if the vocabulary must also be independent of any particular natural language" (Vauquois and Boitet, *ibid.)*.

In the years since CETA was designed, a large body of knowledge has been acquired in the area of representing models of real-world entities in the computer. And even though the task still remains difficult it is more feasible using the modem knowledge representation languages, advanced knowledge acquisition interfaces with built-in consistency and validity checks, suites of programs for processing machine-readable human-oriented dictionaries and encyclopedias, etc. With respect to the choice of names for primitives (the "vocabulary" of Vauquois and Boitet), different knowledge-based systems choose different approaches (e.g., in KBMT-89 the primitives have the status of elements in an artificial language, while in the PREMO system

(Slator and Wilks (1989)) English word senses are used).

Another typical criticism of meaning-based MT, expressed as a criticism of the interlingua approach, concerns the process of TL text generation. Vauquois and Boitet *(ibid.)* write: "The absence of surface-level information makes it impossible to use contrastive knowledge of two languages to guide the choice between several possible paraphrases at generation time." This opinion is seconded by Warwick (1987:28): "One major difficulty with the interlingual approach — aside from the complexity of defining such an abstract model — was that language-specific attributes necessary for defining translation equivalents on the lexical and structural level were neutralised in the interlingual representation, thereby complicating the task of generation considerably."

In a typical transfer system TL generation usually is concerned only with the syntactic part of the process. Text planning and lexical selection are both avoided, the former by uniformly translating every SL sentence by a sentence in the TL,[7] the latter by substituting TL lexical units through bilingual dictionaries. In fact, in early versions of transfer systems generation was little more than a left-to-right scanning and writing out of the terminal elements in a transfer phrase structure tree.

As long as lexical ambiguity is not treated in an MT system, the traditional absence of real lexical selection mechanism is justified simply because there isn't any choice — a single translation variant is suggested for every SL lexical unit. If a more sophisticated variety of the transfer approach can incorporate lexical ambiguity resolution while continuing to use TL as the language for representing the meaning of SL lexical units, then lexical selection in generation may continue to be a non-problem for approaches which use contrastive lexical knowledge. The crux of the matter is, however, still on the analysis side. By using a metalanguage with a higher expressive power than a natural language (we are talking about expressive power for computer programs, not humans!) a meaning-oriented MT system can allow lexical selection in generation to be performed at the level of sufficiently fine-grain semantic features, not monolithic lexical units.

This allows one to smooth out many of cross-linguistic incompatibilities, such as problems of inexpressibility of certain concepts in single-word lexical units in some languages. Multiple examples of such phenomena can be found in MT and general linguistic literature, many of them dealing with translation of kinship terms.[8]

Yet another objection to the interlingua approach to MT is based on "practical" considerations. Bennett and Slocum *(op. cit.)* contend that "since it is not likely that any NLP system will in the foreseeable future become capable of handling unrestricted input — even in the technical area(s) for which it might be designed

[7] This is a very good approximation of a general translation rule, but still constrains the expressive power of a generator. The ability and license to break SL sentences into several TL sentences or combine several of them into one is, as above, a powerful weapon in the hands of a human translator.

[8] One of the latest contributions is Amano (1989) in which it is suggested that Japanese has two lexical units corresponding to the English word *aunt,* one referring to an older sister of a parent and another referring to a younger sister. (In fact, the two words are phonetically the same, though different Kanji characters are used to represent them.) Note that Amano uses this example to support his opinion that direct correspondences between languages alleviate the problem of lexical gaps of this sort. Indeed, his criticism of the interlingua approach includes the statement that, in the cases like the above, use of a descriptive phrase like "father's younger sister" constitutes explanation rather than translation. Following this logic, real translation, then, will necessarily involve either a meaning loss or a potential error in translation. Indeed, for translation from Japanese into English, if the "explanation" mode is to be avoided, both the Japanese lexical items will have to be rendered as "aunt" in English. This is meaning loss. Establishing correct correspondence in the opposite direction will be utterly impossible without extra knowledge (the relative age of the person in question and one of her brothers or sisters) — either using bilingual correlations or using the interlingua method. The difference is that a typical transfer MT system does not have a mechanism to support such an inference even if this knowledge is in principle available, whereas interlingual systems are in principle designed with such problems in mind.

— it is clear that a 'fail-soft' technique is necessary. It is not obvious that such is possible in a system based solely on a pivot language."

"Fail-softness" is a worthy goal for a software system. However, this concept is invoked in the MT literature usually and only to stress a theoretical point, as in the passage quoted just above. In practice, neither transfer-based nor interlingua-based systems have at present a good means of dealing with unexpected or ill-formed input. Nothing in knowledge-based MT *per se* precludes the design and implementation of architectures and algorithms facilitating fail-softness. Just as in transfer systems a target lexical unit can be picked at random (or based on probabilistic judgments, which amounts to the same thing) when no disambiguation is possible, so in interlingua systems some decisions could be made based on similarly weak heuristics. The above criticism is, thus, a non-criticism. It probably stems from the observation that such weak heuristics are seldom used or discussed in meaning-oriented projects because these projects are typically research-oriented rather than devoted to building production system prototypes. However, if such a prototype is built using a meaning-oriented approach the objective of fail-softness can be achieved equally well.

6. How Formal Must Meaning Representation Be?

It is widely supposed that machine translation requires at ground a fully interpreted logical calculus, that a meaning-based approach cannot be presented with such formal rigor and hence that meaning-based MT cannot succeed. This argument may be understood as demanding formal proofs of the correctness of translated meaning representations. Without such proofs, it is supposed, there is no guarantee that a translation will be free of contradiction or that the same meanings will be always represented similarly.

The formalist approach to machine translation is heir to Montague's view that there is or should be no distinction in principle between natural and formal languages. But even if Montague, thus glossed, were correct, it would not follow that uniquely *formal* representations are necessary for the task of machine translation. That is to say. with Wilks (1989:3),

> ... we do need representations (as opposed to the current trend of connectionism...), but their form, if interpretable, is largely arbitrary, and we may be confident it has little relation to logic. I shall restate the view that the key contribution of AI in unravelling how such complex tasks as "understanding" might be simulated by a machine lies not in representations at all but in particular kinds of procedures ... It would be the most extraordinary coincidence, cultural, evolutionary, and intellectual, if what was needed for the computational task should turn out to be formal logic, a structure derived for something else entirely. Although, it must be admitted, strange coincidences have been known in the history of science.

The demand for proofs that a target language text will contain no contradiction is of course a demand that cannot be met. But, fortunately, the problem of avoiding contradiction — in machine translation in particular and natural language processing in general — is an empirical issue and not clearly delimited by formalist claims and purported requirements. That is to say, while it might be nice to be able to offer such proofs, it would be a grievous error to abandon any enterprise unable to provide a formal proof of its future success. Indeed, the formalist gambit has been tried against any number of sciences, including physics, and has come up short.

It is perhaps worthwhile to point to Quine's (1960) admission that the only things that can be radically

translated are the logical connectives. It is not clear how one would press the point, but one might confront formalist demands by suggesting that if the connectives can be deterministically translated, then the (formal) avoidance of contradiction will not be quite so difficult as proposed. At any rate, the intuitions underlying 'not,' 'and,' 'or' and so forth are indisputably common and accessible to natural-language users in the absence of any sort of formalism. If they can be formalized, so much the better for logic; but on what grounds is this formalization required for natural-language *understanding?*

The formalist claim is sometimes made by criticizing uninterpreted formalisms. The elements from which our representations are built are "interpreted" in terms of an empirically constructed domain model rather than through an axiomatically defined set of possible worlds or well-formed formulae in a logical system. To be sure, one must avoid over-facile appeals to future research and empirical criteria as a hedge against formalist strictures. Nonetheless, such a line can productively be deployed against the claim that meaning-based MT cannot ensure that same meanings will get the same translations. If sameness of intralingual meaning is in fact preserved in translation — as corroborated by the judgments of bi- or multi-lingual humans, say — then this should be regarded as *evidence* in favor of the meaning-based approach. It would be folly indeed to disregard such evidence in the absence of a formal proof of the possibility of such evidence!

7. Extractability of Meaning

It is argued that it is impossible to ensure that the meaning can *actually be extracted* from the source-language text and rendered in the representation language. As stated above, the present state of the art does not allow a completely automatic disambiguation and representation of all the semantic and pragmatic phenomena. This is especially true for systems like those coming out of the KBMT project at the Center for Machine Translation, in which the expected results of analysis are very detailed.

Hutchins summarizes the scene as follows (1987:49): "In semantic analysis there has been successful treatment of homography and syntactic ambiguity; and there have been successful implementations of case frames, of semantic features, of distributional semantic information, and recently of Montague semantics; but, nevertheless, the profounder problems of interlingual semantic analysis have proved elusive." These "profounder" problems presumably include treatment of reference (including ellipsis), abductive inference-making on the basis of world knowledge, speaker attitudes, indirect speech acts, stylistic factors, etc. We are making inroads into these and other difficult areas. In the meantime, the reliance on the concept of microtheories, the continued work on the acquisition of domain models and the use of an interactive augmentor (a program which supports the interactive editing functionality to treat those types of meaning which cannot be treated automatically within the current state of the art) make meaning-based systems feasible. This has already been demonstrated at the research level.

One of our tasks is to demonstrate the utility of this approach through a production system prototype. As we mentioned above, the role of the augmentor will progressively diminish as our research on meaning extraction progresses. But it is strange to doubt *that* it is progressing.

8. Translation and Paraphrasing

Hutchins (1987:49) claims that in meaning-oriented MT systems " .. .the abstractness of 'content' representations result[s] in losses of information about 'surface' structures of texts" and from this he concludes

that "versions produced by AI methods are not translations but rather paraphrases."

This opinion relies too much on the formulation of translation as a relation among texts, not among textual meanings (cf. the similar definitions in Johnson *et al.* (1985) and Arnold and des Tombe (1987) as quoted above). If we agree that the invariant in translation is meaning, then translation becomes a paraphrase, only a special one; in this type of paraphrase the lexical, grammatical and prosodic means of a different language are used (see, e.g., Whitelock (1989) for a similar argument).

In fact, the "paraphrasing as translation" argument is a facet of a more general question: Does one need to treat the *form* of the input text during translation? This question naturally arises because the dichotomy of substance and form has been a central point of discussion in such fields as semiotics and art theory and history. By 'form of text' we refer a number of diverse phenomena: the syntactic structure of the input sentences; its phonetic and prosodic properties, such as alliteration, meter, rhyme, etc.; and the layout of a printed page, which can include diagrams, formulas, pictures, examples and other highlighted material, special fonts and so forth.

The layout of the text on a page is a feature independent of text meaning but influences the overall impact of the text. It can be called on to carry an esthetic message (as, for instance, in Apollinaire's poems or Lewis Carroll's tale written in the *form* of a tail in *Alice in Wonderland.* In expository, esthetically neutral text, which is the type of text which is machine-translatable, it is sometimes desirable to preserve page layouts in translation (especially, for pages with diagrams, illustrations, etc.), as, for instance, in the case of multilingual equipment manuals.

It is clearly difficult to preserve phonetic characteristics of the source text in the target text, and not only for computers. We will therefore not expect to deal with these issues. However, the use of special fonts (e.g., italics) carries a meaning which has to be recreated in the target text. Sometimes lexical units from languages other than the main language of the text are highlighted. These should be recognized as material not to be translated but rather reproduced "as is" in the target text. However, in some cases italics are used for purposes of marking sentential stress (e.g., "I do not want to see *any* of them"), and in such cases this meaning should be represented and later re-created in the target language using its own means of expressing sentential stress.

Outside the field of artistic texts — poetry and fiction — preservation of the *syntactic form* of the source text in translation is completely superfluous because the meaning and use of, say, passive voice constructions in a source and a target language should not necessarily be identical. Direct structural correspondences between certain pairs of languages can be exploited in MT systems of a particular type, but they should be treated as idiosyncratic occasions rather than phenomena that occur as a rule and should, therefore, be preserved in translation. From the point of view of quality of expository text translation, it is immaterial whether the syntax of the target sentences is similar to that of the source sentences.

To summarize, there is no reason to aspire to translate the form of the input text. However, if an MT system does not possess sufficient knowledge to analyze SL texts deeply enough to allow understanding sufficient for realization of corresponding TL texts, it may rely on preserving the syntax of the source text in the target text as a very crude decision-making heuristic, regularly violated in a large number of cases.

9. Conclusion

While we have so far emphasized the key points of difference between the two main MT paradigms, it will be productive to conclude with mention of the positions which seem to be held jointly by all MT system developers. These platforms for agreement seem to include the following:

- Translation is a relation between texts in the source and target languages, such that the invariant between them is meaning. In other words, translation is rendering a set of meanings realized in a source language using the realization means of a target language.
- MT deals with expository texts, where the artistic considerations do not play an important role.
- Meanings in such texts are, in practical terms, completely expressible in all relevant source and target languages.
- Fully automated MT is not feasible at present, but
- The main research direction is toward full automation.

Additionally, here are some positions are held by researchers in meaning-oriented MT but are not emphasized by other MT workers:

- SL ambiguity resolution is the main technical goal to be achieved by MT systems.
- Paradigmatic and other design considerations must crucially take into account the above requirement.

Interlingual MT systems tend to favor the meaning-based approach, while transfer systems tend to render meaning without the added requirement of representing it. Theoretically, meaning-oriented MT is not restricted to the interlingua paradigm. One can in principle incorporate meaning analysis into the transfer approach. However, in practice, as such attempts proliferate, it will become clear that the interlingua paradigm is more convenient for the support of the analysis of meaning. We also believe that the amount and complexity of knowledge acquisition for interlingual MT systems is at worst roughly equal to that which would have to be mastered for meaning-oriented transfer MT. At best, the acquisition component of an interlingua approach will be more compact and well-organized.

This paper has dealt exclusively with conceptual arguments. There are a number of practical — technological and methodological — issues relating to the differences among MT approaches. We plan to discuss them in a separate paper.

REFERENCES

Amano, S. 1989. On Interlingua Approaches — from the Point of View of *traduttori traditori.* Presented at the MT Workshop "Into the 90s." Manchester.

Arnold, D. and L. des Tombe. 1987. Basic Theory and Methodology in Eurotra. In: S. Nirenburg, ed., *Machine Translation: Theoretical and Methodological Issues.* Cambridge: Cambridge University Press, 114-135.

Arnold, D., S. Krauwer, L. des Tombe and L. Sadler. 1987. "Relaxed" Compositionality in MT. Proceedings of the Second International Conference on Theoretical and methodological Issues in Machine Translation. Pittsburgh, June 1988.

Bar Hillel, Y. 1960. The Present Status of Automatic Translation of Languages. In: F.L. Alt (ed.) **Advances in Computers** (Volume 1). New York: Academic Press, 91-163.

Ben An, D. M. Rimon and D. Berry. 1988. Translational Ambiguity Rephrased. Proceedings of the Second International Conference on Theoretical and methodological Issues in Machine Translation. Pittsburgh, June 1988.

Bennett, W.S. and J. Slocum. 1985. The LRC Machine Translation System. *Computational Linguistics,* 11: 111-121.

Brown, P., J. Cocke, S. Della Pietra, V. Della Pietra, F. Jelinek, R. Mercer and P. Roossin. A Statistical Approach to French / English Translation. Proceedings of the Second International Conference on Theoretical and methodological Issues in Machine Translation. Pittsburgh, June 1988.

Durand, J, P. Bennett, V. Allegranza, F. van Eynde, L. Humphreys, P. Schmidt, E. Steiner. The Eurotra Linguistic Specifications: An Overview. To appear in *Machine Translation.*

Goodman, K. and S. Nirenburg. 1989. KBMT-89. CMU CMT Technical Report.

Goodman, K. and S. Nirenburg. 1990. To Save the Semantic Phenomena: Machine Translation and Interlingua Texts. Proceedings of the Fifth International Conference on Computers and Philosophy. Stanford, August.

Hutchins, W.J. 1986. *Machine Translation: Past, Present, Future.* Chichester Ellis Horwood.

Hutchins, W.J. 1987. Prospects in Machine Translation. Proceedings of MT Summit. Japan.

Isabelle, P. and L. Bourbeau. 1985. TAUM-AVIATION: Its Technical Features and Some Experimental Results. *Computational Linguistics,* 11: 18-27.

Johnson, R., M. King and L. des Tombe. 1985. EUROTRA: A Multilingual System under Development. *Computational Linguistics,* 11: 155:169.

Katz, J.J. 1988. "The Refutation of Indeterminacy." *The Journal of Philosophy* 85,227-252.

Kay, M. 1973. *The MIND System.* New York: Algorithmic Press.

Kay, M. 1989. Talk at MT Summit II. Munich. August.

King, M. 1981. Design Characteristics of a Machine Translation System. Proceedings of IJCAI-81. 43-46.

Landsbergen, J. 1989. The Rosetta Project. Proceedings of MT Summit II. Munich. 82-87.

Nagao, M. 1989. *Machine Translation: How Far Can It Go?* Oxford: Oxford University Press.

Nirenburg, S. and V. Raskin. 1986. A Metric for Computational Analysis of Meaning: Toward an Applied Theory of Linguistic Semantics. Proceedings of COLING-86, Bonn, 338-340.

Pericliev, V. 1984. Handling Syntactic Ambiguity in Machine Translation. Proceedings of COLING-84, 521-524.

Quine, W.V.O. 1960. *Word and Object.* Cambridge, Mass.: The MIT Press.

de Roeck, A. 1987. Linguistic Theory and Early Machine Translation. In M. King, ed.. *Machine Translation Today*. Edinburgh: Edinburgh University Press, 38-57.

Schneider, T. (1989). The METAL System. Status 1989. Proceedings of MT Summit II. 128-136.

Slator, B. and Y. Wilks. 1989. PREMO: Parsing by Conspicuous Lexical Consumption. Proceedings on the International Workshop on Parsing Technologies. Pittsburgh, August, 401-413.

Uchida, H. 1989. ATLAS. Proceedings of MT Summit II. Munich. 152-157.

Vasconcellos, M. and M. Leon. 1985. SPANAM and ENGSPAN: Machine Translation at the Pan American Health Organization. *Computational Linguistics,* 11: 122-136.

Vauquois, B. and C. Boitet. 1985. Automated Translation at Grenoble University. *Computational Linguistics,* 11: 28-36.

Warwick, S. 1987. An Overview of post-ALPAC Developments. In: M. King (ed.) Machine Translation Today. Edinburgh University Press. 22-37.

Whitelock, P. 1989. Why Transfer and Interlingua approaches to MT are Both Wrong: A Position Paper. Presented at MT Workshop "Into the 90s." Manchester.

Wilks, Y. 1989. Form and Content in Semantics. Manuscript, Computing Research Laboratory, New Mexico State University, Las Cruces.

Wilks, Y. 1988. Philosophy of Language and Artificial Intelligence, Technical Memorandum MCCS-88-132, Computing Research Laboratory, New Mexico State University.