Meta
Journal des traducteurs
Translators' Journal

MΕΤΑ

Basic Concepts of MT

Wolfram Wilss

Volume 38, numéro 3, septembre 1993

URI : https://id.erudit.org/iderudit/004608ar
DOI : https://doi.org/10.7202/004608ar

Aller au sommaire du numéro

Éditeur(s)

Les Presses de l'Université de Montréal

ISSN

0026-0452 (imprimé)
1492-1421 (numérique)

Découvrir la revue

Citer cet article

Wilss, W. (1993). Basic Concepts of MT. *Meta*, *38*(3), 403–413. https://doi.org/10.7202/004608ar

Résumé de l'article

Malgré les progrès réalisés dans l'élaboration des algorithmes de parsage, les systèmes de traduction automatiques, même les plus sophistiqués, se heurtent encore à de nombreux problèmes dont ceux de la complexité syntaxique, de l'ambiguïté lexicale et de l'analyse interphrastique. La compréhension de la corrélation entre les diverses unités syntagmatiques dépend des connaissances linguistiques, extralinguistiques et contextuelles dont la combinaison permet justement au traducteur humain de résoudre de manière quasi automatique les problèmes d'ambiguïté. La mise sur pied d'un système totalement indépendant de l'action humaine relevant aujourd'hui de l'utopie, la recherche en traduction automatique devrait davantage s'orienter vers la conception de micro-systèmes utilisables dans des domaines précis.

BASIC CONCEPTS OF MT

WOLFRAM WILSS
Universität des Saarlandes, Saarbrücken, Allemagne

Résumé

Malgré les progrès réalisés dans l'élaboration des algorithmes de parsage, les systèmes de traduction automatiques, même les plus sophistiqués, se heurtent encore à de nombreux problèmes dont ceux de la complexité syntaxique, de l'ambiguïté lexicale et de l'analyse interphrastique. La compréhension de la corrélation entre les diverses unités syntagmatiques dépend des connaissances linguistiques, extralinguistiques et contextuelles dont la combinaison permet justement au traducteur humain de résoudre de manière quasi automatique les problèmes d'ambiguïté. La mise sur pied d'un système totalement indépendant de l'action humaine relevant aujourd'hui de l'utopie, la recherche en traduction automatique devrait davantage s'orienter vers la conception de micro-systèmes utilisables dans des domaines précis.

Machine translation (MT) has been a fascinating subject of linguistic, terminological, psycholinguistic and computational research in the second half of this century. The fascination stems from the fact that, according to Maurice Gross, "mechanical translation is perhaps the first attempt to apply computers to the simulation of a (nonnumerical) human activity." (Lehrberger and Bourbeau 1988: XI) The degree of interest, failure and success for MT has varied over the last 25 or 50 years, but after a first phase of euphoria in the 50s and a phase of disillusionment in the 60s, MT now finally seems to have established itself as a field of interdisciplinary research in which the newly emerging field of Artificial Intelligence (AI) plays a dominant role.

The reason for the rapprochement between MT and AI research is due to the fact that the obstacle to efficient MT with a (commercially) reasonable input-output relation is mainly linguistic. In different words: MT is a software rather than a hardware problem (although hardware problems are by no means totally negligible). Linguistic performance — at any rate in MT-relevant texts — is based on a specific set of source-language (SL) and target-language (TL) rules and configurations, and it is obviously extremely difficult to harmonize these sets of rules and configurations in such a way (and to such an extent) that computer experts are in a position to write programmes which take sufficient account of the intricacies which are characteristic of situationally embedded natural language activities.

The principal impediment to MT is that the computer, or, for that matter, the computer programme, does not "understand" a text to be translated, at least not in the human sense of the word. We all know that in understanding a linguistic utterance-in-context, three (interactive) levels of comprehension must be distinguished, syntactic, lexical, and pragmatic understanding (Enkvist 1987, in Hüllen-FS). Of these three levels or areas of understanding, computer programmes can handle syntactic properties to some extent, semantic properties to a far lesser extent, and pragmatic properties which are related to world knowledge or "possible worlds" to no extent.

The challenge for MT research is therefore to develop MT programmes which are able to simulate the capability to the human translator to extract the meaning of the words and sentences that form a text. Here AI research comes in, because it is one of the chief targets of AI experts to conceive of formalized strategies which enable the computer to behave in a quasi-human manner. The ultimate goal of MT-related AI research is to crack the blackbox of the human translator and to access in a formalized manner his purposive, declarative and procedural knowledge.

In a way, this is, of course, also the goal of human translation research, and for this purpose it has developed (at least) three alternative approaches:

1. the so-called aloud-thinking method;
2. the retrospective aloud-thinking method;
3. the reconstruction of translation procedures, *e.g.* in the context of error analysis.

For MT research, the formalization of knowledge is imperative; hence the importance of knowledge-based MT systems as they are now being developed in the USA and in other places. That the formalization of knowledge is the grass-roots problem of efficient MT has been recognized by MT research for a long time, but it has been unable so far to find the necessary ways and means to achieve this immensely complicated target.

Apart from the aforementioned three levels of knowledge, syntactic, semantic, and pragmatic, in MT research we must distinguish two additional types of knowledge, implicit and explicit knowledge. Implicit knowledge, also known as "tacit knowledge," is the type of knowledge which the human translator controls more or less intuitively or on the basis of contextual information (*e.g.* by reference to information in the preceding sentences or paragraphs). Good specimens of implicit knowledge are anaphoric relation or the (superficially unexpressed) semantic dependency relations between the various components of compounds.

In contrast to the human translator, implicit knowledge is a textual feature that has turned out to be a perpetual and self-perpetuating plague for any fully automatic MT system (in the sense of Yehoshua Bar Hillel's concept of "Fully automatic high-quality translation" (FAHQT)). This is particularly true of supra-sentential relations, because at the moment, at least to my knowledge, there exists no MT system, not even in a rudimentary fashion, with supra-sentential parsing capacity to analyse such relations. Intra-sentential analysis is the utmost which any MT system can handle, and even this is possible only within the confines of fairly narrowly limited syntactic complexity and domain-specificity (for the latter aspect, see below). A somewhat lesser, but still largely unsolved issue is the analysis of compound constituents. An operational (commercially useful) MT system should be able to analyse at least compounds with a fairly simple (and repetitive) structure, for the obvious reason that such a compound analysis module would eliminate "the need to place in the dictionary all compound words or all groups containing a compound word. If a system does not have rules for dealing with the principal mechanisms of compound word construction, updating the dictionary runs the risk of being a constant, never-ending task." (Lehrberger and Bourbeau 1987, 1988)

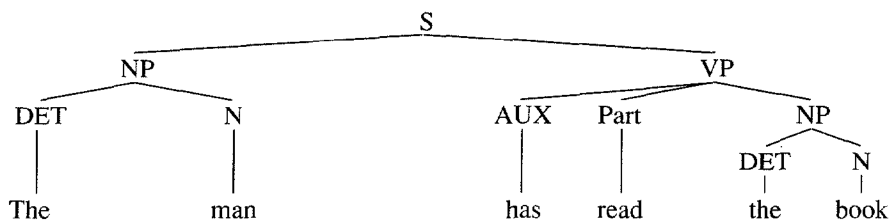
If we take this quotation to its logical conclusion, it becomes apparent that the quality of an MT system principally depends on two interrelated factors:

1. the degree of sophistication in the design of an MT system;
2. the ability of an MT system to deal with problems within the framework of what is nowadays called the "depth-analysis approach."

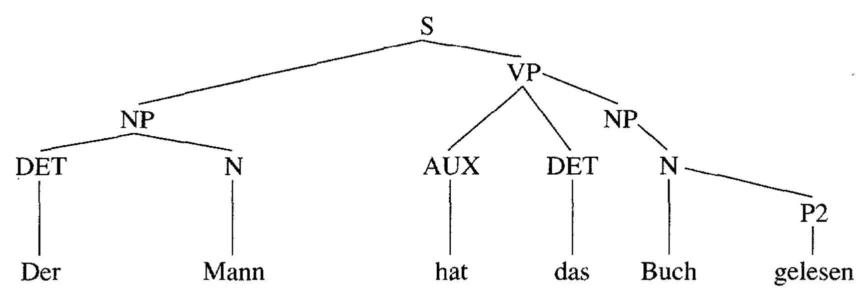
According to the seminal work by Lehrberger/Bourbeau, MT research tends to distinguish four sophistication levels (which, at the same time, represent four logically subsequent MT generations):

1. word-for-word substitution (one-to-one correspondences);
2. machine-aided human translation (MAHT);
3. human-aided machine translation (HAMT);
4. fully automatic machine translation (FAMT).

As the history of MT research has abundantly demonstrated, word-for-word substitution, which amounts to a rigorous form of interlinear version known from the Middle Ages, is, at any rate commercially, of no value, because even an expert in a certain domain will have great, if not unsurmountable, difficulties in extracting the meaning of a document which is partly or fully unintelligible because of the surface distortions of the target-language text (TLT). A possible exception are structurally very simple sentences, containing, say, a subject phrase consisting of an article plus a noun, such as "the man," a verbal phrase consisting of an auxiliary and a past participle, such as "has read," and an object phrase consisting of an article plus a noun, such as "the book." Such a configuration can be illustrated by the following tree structure:



Even those people who have only a very limited expertise in translation know, of course, that syntactic configurations of this type are rare in natural language text production, and, to make things more complicated, even such simple structures require some word-order shifting in going, *e.g.*, from English to German:



The next higher level of sophistication in automation is MAHT: MAHT is basically a human translation effort with only a limited aid from the computer, *e.g.* by providing the translator with an alphabetically organized — or, even better, a textually organised — list of TL equivalents. In other words: the machine component of an MAHT system simply consists of a word processor with the capability of looking up TL terminological correspondences and, possibly, with the provision for sorting these equivalents according to the needs of the translator, as he moves along the text-to-be-translated. This procedure is, according to evaluations, *e.g.* by the *Bundessprachenamt* in Hürth, faster than cumbersome thumbing through a hard-copy (conventional) dictionary, but the responsibility of translating the respective text rests entirely with the translator. Nevertheless, MAHT should not be disposed of lightly (in contrast to word-for-word substitution systems which are out anyway), because MAHT systems may contain a gadget known as KWIC facility. KWIC stands for "key-word-in-context" and it is a device to specify certain words in certain textual environments, thus helping the machine to locate specific lexical items in particular linguistic or domain-specific context and thus to facilitate the disambiguation of homographs.

Conversely, we speak of HAMT, if the computer has a programme at its disposal enabling man-machine interaction in which the computer performs (basic) translation activities, with the human translator providing additional information to "fill out" the computer programme and to prevent it from producing unintelligible garbage. While the computer supplies a raw output, it is the job of the human translator to complete the MT programme by doing more or less refined post-editing work (Vasconcellos 1986). In addition to post-editing, the human translator may also engage in pre-editing activities, thereby paving the way for relatively smooth and economical MT operations. A third and fourth alternative in HAMT have been described by Lehrberger and Bourbeau: "The machine may pause in mid-sentence to query the operator and then resume its processing of the remainder of the sentence, or it may make more than one pass through the whole sentence, with the operator inserting the appropriate information between passes." (1988: 7)

Nowadays, HAMT is widely used commercially, because fully automatic high-quality translation is to all intents and purposes utopian, and even if it were a realistic target for the near future in MT research, the development of pertinent programmes would be so forbiddingly expensive and time-consuming that conceivably nobody would be interested in commercial exploitation.

The intricacies of fully automatic high-quality translation are mainly due to three reasons:

1. syntactic structures beyond a certain complexity threshold have, as indicated, proven extremely cumbersome to parse automatically;
2. the disambiguation of lexical ambiguity, *i.e.* the contextual disambiguation of words with multiple meanings, besets MT research with many unsolved, and possibly unsolvable, problems;
3. the marking off of intra-sentential phrasal constituents is, as, among other things, the protracted discussion of so-called gardenpath sentences shows (Wilss 1988, Kap. XI), very difficult to achieve. Here are three examples, one for syntactic complexity, one for lexical ambiguity, and one for intra-sentential constituent analysis:

(1a) ...the girl who talked to the teacher about the problem...
(1b) ...the teacher whom the girl talked to about the problem...
(1c) ...the problem that the girl talked to the teacher about...

This example, which I have borrowed from Halle *et al.* (1978: 132), is obviously very tricky for the machine to analyse. The reason for this is that the prepositional phrase "about the problem" is marked by disconstituency, with "that the girl talked to the teacher" contextually intervening between "the problem" and the postponed preposition "about," thus making it virtually impossible for the machine to determine correctly the dependency relations within the prepositional syntagme "about the problem."

(2) She cannot bear children

In this sentence, the crucial item is "bear," meaning either that she has no way of getting along with children, or that she is physically unable to give birth to children. Disambiguation requires contextual (situational) information which is beyond the range of MT parsing algorithms.

(3a) When turning over control operators will switch to zero.

This sentence permits two, theoretically even three, different syntactic (and semantic) readings:

(3b) When turning **over** control, operators will switch to zero
(3c) When **turning** over control, operators will switch to zero
(3d) When turning **over**, control operators will switch to zero

Of course, every reading entails a specific (German) TL meaning:

(4a) *Bei der Übergabe des Kontrolldienstes (Ablösung vom Kontrolldienst) ist auf Null zu schalten*
(4b) *Beim Wenden über dem Kontrollpunkt ist auf Null zu schalten*
(4c) *Beim Wenden schaltet das Kontrollpersonal auf Null*

I am unaware of any commercial MT system with enough power for syntactic and/or semantic parsing of the above examples, let alone for an adequate transfer into a TL. Hence, I take a pretty dim view of fully automatic MT, at least for texts with uncontrolled syntax and uncontrolled lexicon. Things look different for syntactically and semantically rigorously restricted MT systems, as the Canadian "METEO"-project shows (translation of weather forecasts from English into French), but more ambitious projects with less controlled syntax and bigger lexicon are for a machine unmanageable, as the failure of the Canadian "Aviation"-project has demonstrated.

If we compare MT systems of the MAHT or the HAMT type, it becomes apparent that the two systems represent in principle different degrees of MT sophistication, but it is in practice probably difficult to draw a hard-and-fast dividing line between them, because in the last analysis it is in both cases the person of the human translator who has the ultimate control over the MT operations. The prevalent drawback of the two systems is in my opinion that the indispensability of translator (or bilingual expert) interference may slow down the whole MT process to the point where the volume of texts handled in an MT system either of the MAHT or the HAMT type in a certain amount of time may not be significantly greater (or may be even smaller) than that of the human translator doing his job without computer assistance. The point is that machine output will in most cases place greater demand on the post-editor than would human output place on the revisor. A wayout may be "quick-fix" post-editing setting as its target only minimal understandability and readibility of MT production.

In contrast to MAHT and HAMT the principle of fully automatic MT (FAMT) is human non-intervention between SL input and TL output. In other words: FAMT is conceived of as a non-interactive system. According to Lehrberger and Bourbeau, there is only one FAMT system (METEO), "in which the machine itself decides which of the sentences submitted to it are to be revised, all others being translated and considered suitable as finished text ready for use." (1988: 8) But in the strict sense of the word, METEO does not wholly fulfill FAMT requirements.

If we now switch over to the aspect of depth of analysis, we may as well start off our discussion with the by now rather trivial statement that we must make a principal distinction between human and machine parsing behaviour. For the human translator, there is no fixed depth level; rather, depth of analysis is a largely individual phenomenon, because individual translators can adjust the SL environment to the TL text environment in different ways. Whereas the general processual components in translational behaviour are more or less universal, their concrete manifestations in the build-up of textually appropriate analysis and transfer strategies, methods and techniques are likely to vary from one person to another, according to the impact and range of adotive skills (problem-solving routines). What adaptive skills consist of may differ, at least to some degree, both across translators and across the cognitive demands of individual translational situations. So adaptive performance is not quite the same thing for different translators and in different settings. The awareness of the need for compensation may be the same, but how compensation is achieved largely depends on the adaptive skills the translator is able to capitalize upon.

In contrast to the human translator, this sort of flexibility is non-existent in MT systems. The machine operates on a pre-determined programme which is organized according to the rather rigid principles of a flow-chart, allowing virtually no leaps or jumps on its way through its various modules.

Theoretically, we can distinguish three levels of magnitude in depth-of-analysis operations:

1. syntagmatic (analysis of word groups);
2. syntactic (analysis of sentence configurations);
3. textual (macro-analysis).

SYNTAGMATIC ANALYSIS

Syntagmatic depth of analysis is also known as "local analysis," the range of analysis being restricted to word strings which in conventional grammars are called speech parts, such as subject phrase, predicate phrase, object phrase, attributive phrase, etc. The limitations of local analysis become apparent, if we look at two sentences which — together with the main arguments of the commentary — I have taken over from Lehrberger and Bourbeau (1988: 9) (leaving out French equivalents):

(5a) The passenger flight arrival time changes every year;
(6a) The passenger flight arrival time changes will be posted.

Before translating "the" in these sentences, the computer (programme) must find the head of the noun phrase in each case. But with this last-noun-in-a-sequence strategy the computer may run into difficulties when it tries to locate the head noun in (5a): It must consider — the reader forgive me for humanizing the computer — the possibility (feasible or not) that "changes" is a plural noun, although it is, in fact, a verb in (5a) and "time" is the head noun. Furthermore, in both (5a) and (6a), the computer must eliminate, by applying context-adequate inference strategies, the possible identification of "time" as a verb. It is only by examining the entire sentence that the computer can — hopefully — be sure that the whole nominal phrase has been correctly identified. Those who are familiar with the universe of discourse will realize immediately that such configurations are extremely frequent, as *e.g.* any LSP dictionary will testify, and that things may become even trickier by the insertion of embedded phrases which may occur in various degrees of complexity.

The arguments presented just now show that local analysis is O.K. for simple syntactic configurations of the type resulting from syntactic paraphrase operations of the two last-mentioned sentences:

(5b) The arrival time / of passenger flights / changes / every year;
(6b) Changes / in the arrival time / of passenger flights / will be posted.

The amount of syntactic restructuring in MT (say: German) and yet the parsimonious effect on the depth of analysis is considerable, because dependency relations between the word strings of the individual phrasal units are far more transparent than in (5a) and in (6a). This explains the attempts to design MT systems with the already mentioned property of controlled syntax a good example of which is the French project "TITUS."

FULL SENTENCE ANALYSIS

Any MT expert will testify to the fact that holding the full sentence in perspective is an indispensable precondition for efficient analysis. The interdependence between the various syntagmatic units in a sentence requires three types of knowledge, linguistic knowledge, extralinguistic or real world knowledge, and situational knowledge. This is demonstrated by the following four examples:

(7) Flying planes can be dangerous;
(8) Although flying planes can be dangerous, he decided to apply for a job as a pilot;
(9) Flying planes can be dangerous, but it is profitable;
(10) Flying bees can be dangerous.

These four examples hardly need commenting: whereas (7) (a well-known Chomsky example) is syntactically and semantically ambiguous and therefore requires for its disambiguation supra-sentential information, the remaining three examples are semantically self-contained and are therefore, at least theoretically, amenable to MT parsing. A complicating factor in full sentence analysis is the chain-reaction effect which means that, as a rule, local analysis within a sentence cannot be carried out in isolation without affecting other syntagmatic units in one way or another.

TEXTUAL LEVEL OF ANALYSIS

As already indicated, text-level analysis is possible only to the extent that supra-sentential dependency relations are fully explicit (thus, *e.g.*, forbidding anaphoric reference) and requiring a high degree of lexical redundancy or repetitiveness. It is, however, doubtful that this type of text strategy is enforceable on LSP writers. At any rate, automatic text grammars, in contrast to automatic sentence grammars, are non-existent and, to my knowledge, there are no signs of much progress in the foreseeable future. Supra-sentential transitive strategies are so multifarious that it is wishful thinking to hope to work out standardized MT analysis procedures with noticeable probability rating.

After discussing various types and degrees of computerized analysis, let us now turn to the equally important aspect of computerized transfer. There are, roughly speaking, two types of transfer design. An MT system may be characterized either as a direct, one-stage transfer, or as an indirect, multiple-stage transfer. The latter is also known as the "pivot language" transfer or the interlingua type of transfer. The terms "direct" and "indirect" transfer are almost self-explanatory, but nevertheless some illustrative remarks may be in order, especially since these two terms also occur in human translation research. Whereas the meaning of the term "direct" transfer is more or less comparable in human translation and in MT, the concept of "indirect" transfer is different, because the interlingua type of transfer is unknown in human translation, unless we mean by "interlingua" interim transfer stages of the multiple-stage translation introduced in human translation research, to my knowledge, by Voegelin (1954; for details see Wilss 1988, chapter IV and Wilss 1992, chapter IV). In human translation the terms "direct" and "indirect" transfer are — *cum grano salis* — equivalents to "literal" vs "nonliteral" translation.

When we talk of "direct" transfer, we mean that a transfer system, no matter whether human or machine, tries to take the shortest possible route from a SL text to a TL text unit. This type of transfer is characterized basically by two complementary processes:

1. lexical replacement or substitution;
2. syntactic adjustment (rearrangement) or shift of expression.

In order to be operative, a "direct" transfer system requires two components:

1. a bilingual dictionary, hopefully with no multiple-meaning words. This type of dictionary is best represented by LSP glossaries which supply the user with just one TL equivalent for just one SL item, leaving no leeway in the choosing of the correct TL correspondences;
2. a programme for syntactic rearrangement by changing the SL word order according to the rules of the TL, adding or deleting items where appropriate, etc.

Examples for lexical equivalents are:

(11) source language / target language — Ausgangssprache / Zielsprache, direct / indirect transfer — Direkttransfer / indirekter Transfer, one-to-one-correspondence — Eins-zu-Eins-Entsprechung, human / machine translation — menschliche / maschinelle Übersetzung, artificial intelligence — künstliche Intelligenz

Of course, each glossary item should be accompanied by additional information such as grammatical categories (speech parts), subcategories, plus a minimum of semantic (collocational and idiomatic) properties, information on derivational and compositional morphology, etc.

A particularly important aspect of collocational semantics are the so-called selectional restrictions for which Lehrberger and Bourbeau have adduced a very impressive (by no means exceptionally complex) example (1988: 14):

(12) The student enters the classroom
The robber enters the house
The man enters a horse in a race
The young man enters the service
The book-keeper enters an item in the ledger
The lawyer enters an action against the accused
The defendant enters a plea of not guilty
The thought never entered my head

It is obvious that the meaning of "enter" cannot possibly be determined by just looking at the verbal phrase; in all cases the correct meaning depends on the subsequent object phrase (and in some cases also on the preceding subject phrase). In going, *e.g.*, from English to German, instead of just having to look for one comprehensive equivalent of "enter," we must be prepared — and this must be taken care of in the built-up of the machine dictionary entry — to operate with several lexical equivalents (in the worst case with eight). All these arguments show convincingly that only if the computer dictionary is equipped with sufficiently detailed collocational information will the programme be able to turn out the correct translation. This has led Laffling to conceived of a lexicon-driven MT approach. He believes that recourse "to the world knowledge and "deep" understanding is perhaps not as necessary as some Al researches would posit...." (1991: 114) A similar stand is taken by Starosta who has developed the "lexicase" concept. (1988)

Similar selection constraint problems occur in German-English MT (*e.g.* for the verb "einstellen"), because, as valency grammar research has shown, lexical items (*i.e.*

verbs) with a complex relational infrastructure are also typical of the German language. Hence the need to organize the lexicon in such a way as to enable the computer to process sentences on the basis of a recursive algorithm or inference mechanism, taking as much account as possible of predicate-argument structures explaining the role of speech parts (arguments) in connection with the respective verb usage. This requires a lot of subtle linguistic ground work. It is perhaps somewhat exaggerated but in principle true that, as Professor Guenther from Munich University has stated, we might as well need a full doctoral dissertation for documenting the full collocational and idiomatic range of each individual German verb. The same may be true of English verbs.

However, the development of transfer rules in accordance with selectional restrictions is by no means the only problem in the design of an MT system. Closely related (and somewhat overlapping) is the problem (issue) of how to handle idioms or phraseological units in an automatic dictionary. *E.g.*, a literal translation into German of "red tape" would be completely meaningless, and the same would be true of a literal translation of the idiom "to kill two birds with one stone" or "to bark up the wrong tree," because there are no literal equivalents in German. Note that more information about the specific features of idiomatic expressions and equivalence problems is available from the "Europhras"-collection of articles published in 1988 by Gréciano from Straßburg University.

The biggest headache-provoking problem in MT research is, as already mentioned, ambiguity (and the related problem of homography, *i.e.* the occurrence of lexical items in more than one part-of-speech function). The frequency of ambiguous words varies from text to text and may be less prevalent in LSP material than in common-language texts, but ambiguity is nevertheless a principal bane for MT researchers, because no overall solution is in sight. While the human translator is rarely bothered by ambiguity, because disambiguation takes place almost automatically through the combination of linguistic, extra-linguistic, and situational knowledge bases, an MT system, for the simple fact that it lacks an encyclopedic memory, is unable to determine the textually appropriate meaning of a multiple-meaning word by using decision making procedures. Hence, whenever an ambiguous word occurs, the computer must pursue, nonsensical as it may be, every possibility that is theoretically open to it. This is so, because the computer treats lexical items as meaningless symbols or symbol combinations which are concatenated in a sentence by more or less elaborate syntactic rules or rule combinations. This predicament is somewhat compensated for by the printout of all possible semantic interpretations, leaving the correct interpretation to the user of machine-produced translations. This problem can somewhat be alleviated by listing, on the basis of probability ratings, only one (domain-specific) (default) meaning and, as a consequence, by ignoring all other meanings as candidates for lexical sub-entries. Of course, this simplification procedure runs the risk of provoking "failure" or "word missing" signals on the screen or on a hard-copy printout, but this is the sort of bottleneck MT systems designers must live with for the time being.

Another LSP-specific phenomenon of ambiguity is the transformation of English participle construction into finite sentences, as prescribed by German syntactic rules (for a detailed discussion see Wilss 1988, Chapter XI):

(13) Having finished his job, he went home
(14a) *Nachdem er seine Arbeit beendet hatte, ging er nach Hause*
(14b) *Da er seine Arbeit beendet hatte, ging er nach Hause*

The participle construction is marked by the absence of a subject phrase making necessary what in modern linguistic has been called "subject raising." As a sort of syntactic compensation, there is the "ing"-suffix. Its usefulness for syntactic analysis is, however, almost completely neutralized, because it is syntactically polyvalent.

To make things worse, there are not only syntactic identification problems. Participle constructions can, as the two examples (14a) and (14b) show, express different semantic modalities which are not represented on the surface level. Syntactic relations and semantic relations are not coextensional. Seen from this angle, participle constructions may be regarded as an analogy to compound words which Jespersen has characterized as follows: "The merit of compounds lies in their conciseness... Compounds express a relation between two objects or notions but say nothing of the way in which the relation is to be understood. That must be inferred from the context or otherwise..." (1954: 137)

On the basis of the last discussion, we can make three statements:

1. sentence meaning does not always constitute itself as a set of explicit linguistic features;
2. linguistic signs and sign combinations differ as to their — metaphorically speaking — contextual "radiating power;"
3. translation is a psycholinguistic operation presupposing self-supporting intelligence and the ability to perform semantic interpretation of the SL text guaranteeing a coherent transfer into the TL.

One final point: the future of MT most probably lies in the area of HAMT rather than in the area of FAMT. If HAMT research wants to achieve more than makes shift solutions, it must take steps to improve the two main process components, substitution and restructuring. This is necessary to cut down post-editing (or pre-editing) to the absolute minimum and thus to avoid commercially uninteresting duplications of effort in interactive systems.

Now, what are the strategies to achieve this doubtless highly ambitious target? In my view, there are two approaches, one predominantly theoretical, one predominantly pragmatic:

1. the theoretical approach has to be seen within the spectre of recent trends in linguistics which Hudson (1990: 3) has telescoped into the term "word grammar" and listed as follows:
   "a) lexicalism: the tendency to shift explanations from facts about constructions to facts about words;
   b) wholism: the tendency to minimize the distinction between the lexicon and the rest of the grammar;
   c) trans-constructionism: the tendency to reduce the number of rules that are specific to just one construction;
   d) poly-constructionism: the tendency to increase the number of particular constructions that are recognized in a grammar;
   e) relationism: the tendency to refer explicitly to grammatical relations, and even to treat these, rather than constituent structure, as primary;
   f) mono-stratalism: the tendency to reject the transformational idea that a sentence's syntactic structure cannot be shown in a single structural representation;
   g) cognitivism: the tendency to emphasize the similarities and continuities between linguistic and non-linguistic knowledge;
   h) implementationism: the tendency to implement grammars in terms of computer programs."

2. The pragmatic approach is in an exemplary fashion represented by the concept of "sublanguage" which has primarily been developed by Kittredge and Lehrberger (1982), but recently this idea has also found an increasing number of supporters in

the USA, in the Soviet Union, and other countries. The basic concept is the domain-specificity of the bulk of LSP material. It starts from the assumption that texts within a given realm of scientific or technological activities are usually characterized by a special terminology and, at least to some extent, also by a restricted syntax. METEO is, of course, a good case in point. Additional stimuli in this direction are widespread user needs demanding domain-dependent systems, the chance to extend the practiced methodology into other fields and to evaluate existing systems with the aim of finding out whether and to what degree such domain-dependent systems are capable of satisfying user requirements.

However, domain-specificity is not tantamount to a final break-through in MT enabling the discharge of human translators en masse. A more realistic prediction is an increase of the use of computers by human translators who continue to perform the actual translation and, parallel to this development, an increase in the number of terminologically-based small-scale MT systems helping the human translator with his job in limited domains. These systems should be sophisticated and at the same time economical enough to relieve the human translator of routine jobs and to make machine-produced pre-translation with a tolerable amount of pre/post-editing feasible. More ambitious targets, such as multi-lingual MT, working with an interlingua component, or extension of MT to other subject-areas outside LSP material, are, in my view, presently not worth considering.

REFERENCES

ENKVIST, N. E. (1987): "More About Text Strategies", W. Lörscher and R. Schulze (Eds.), *Perspectives on Language in Performance. Studies in Linguistics, Literary Criticism, and Language Teaching and Learning*, Festschrift for Werner Hüllen on the Occasion of his sixtieth Birthday, Tübingen, pp. 337-350.

GRÉCIANO G. (Hg) (1989): *Europhras 88. Phraséologie contrastive. Actes du colloque international, Klingenthal-Strasbourg, 12-16 mai 1988*, Collection recherches germaniques n° 2, Strasbourg.

HALLE, M. *et al.* (1978): *Linguistic Theory and Reality*, Cambridge, Mass.

HUDSON, R. (1990): *English Word Grammar*, Oxford / Cambridge, Mass.

JESPERSEN, O. (1954): *Modern English Grammar*, Vol. VI, London / Copenhagen.

KITTREDGE, R. and J. LEHRBERGER (Eds.) (1982): *Sublanguage. Studies of Language in Restricted Semantic Domains*, Berlin / New York.

LAFFLING, J. (1991): *Towards High-Precision Translation Based on Contrastive Textology*, Berlin / New York.

LEHRBERGER, J. and L. BOURBEAU (1988): *Machine Translation. Linguistic Characteristics of MT Systems and General Methodology of Evaluation*, Amsterdam / Philadelphia.

STAROSTA, S. (1988): *The Case for Lexicase. An Outline of Lexicase Grammatical Theory*, London / New York.

VASCONCELLOS, M. (1986): "Functional Considerations in the Postediting of Machine-Translated Output", *Computers and Translation*, 1-1, pp. 21-38.

VOEGELIN, C. F. (1854): "Multiple Stage Translation", *International Journal of American Linguistics*, 20, pp. 271-280.

WILSS, W. (1988): *Kognition und Übersetzen. Zu Theorie und Praxis der menschlichen und der maschinellen Übersetzung*, Tübingen.

WILSS, W. (1992): *Übersetzungsfertigkeit. Annäherungen an einen komplexen übersetzungspraktischen Begriff*, Tübingen.