On the Proper Place of Semantics in Machine Translation*

Margaret King
ISSCO
University of Geneva
e-mail: king@divsun.unige.ch

Preface

This paper owes much, both directly and indirectly, to Don Walker. Indirectly, it concerns a topic which was of great interest to both of us, and which we often discussed. Don helped me to clarify my own ideas and was always both perceptive and practical, a rare combination which made discussion with him constantly illuminating. There is also a direct debt; it was written at a time when I was hors combat from the normal round of life. Don encouraged me to use the opportunity to work out what I really thought without the usual time pressure, and cheered me up in the patches of loneliness. In brief, he lent me some of his own courage.

I am very grateful to Professor Makato Nagao and to the original publishers for allowing me to contribute something which brings back to me strong and good memories of a friendship I valued enormously.

1 Introduction

In this paper I want to ask first what it means to be able to translate between two languages, then use the results to examine the status of a number of semantics based mechanisms frequently proposed for use in machine translation systems. The thrust of the argument will be that it makes no sense to look for semantic or epistemological universals, with the intention of basing translation on language independent abstract entities, but that such mechanisms as semantic features or deep case roles should rather be thought of as engineering tools to be used within a view of translation as essentially a linguistic enterprise.

2 Meaning and Translation

Much of what one believes about translation depends on what one believes about meaning. A once prevalent view of the question in philosophy would have us believe that somehow the 'meanings of words are entities independent of the words themselves', so that for any word or phrase, there is something which could be called the meaning of that word or phrase. Thus, a theory of meaning would consist in trying to say in a general way what kinds of things meanings were, with there perhaps being several different kinds of meanings, depending on the theory. Quine gives a graphic summary of this view at the beginning of *The inscrutability of reference:* 'Uncritical semantics is the myth of a museum in which the exhibits are meanings and the words are labels. To switch languages is to change the labels.'

I shall not, here argue against the 'independent entities' view in any detail: that has been done by many others before, and far better than I could do it. (For example, by Wittgenstein, Ryle, Austin and Quine himself - to name but a few).

What is more interesting for us here is to look at some of the alternatives that have been put forward, and see what the consequences are for a view on translation. Nearly all the alternatives start by denying the assumption that one can name or identify the meaning of a phrase or a word, and by suggesting that we have been misled by the fact that one can ask 'What does x mean?' into believing that that question is the same sort of question as 'Who is Rachel's maths teacher?', which clearly has a coherent answer by naming ('Mr. Dupont') or by identifying ('the man with red hair'). This sounds rather like just saying that the view is wrong and going no further, but if we look at the way we normally talk about meaning, for example, at what is involved in explaining the meaning of a word and at what is pre-supposed by our ways of explaining, it becomes rather clear that we do not appeal to any notion of independent meanings, and thus the charge that a wrong question has been asked gains some support. Both Wittgenstein and Austin approach the question in this way, by asking how we explain what something means. This leads to the suggestion that we substitute for the 'independent meanings' view, the ordinary conception of meaning, what Caton calls 'the everyday concept in daily use.' The most famous statement of this is Wittgenstein's, in the *Philosophical Investigations* 'For a large class of cases – though not for all – in which we employ the word "meaning" it can be defined thus: the meaning of a word is its use in the language.' (1953, 34).

Another way of stating this would be to say that there is no point in looking for the meaning of a word or phrase outside the language in which it is a word or a phrase: in other words, meaning is language internal and critically bound to the particular language. (Actually, it is not totally clear that Wittgenstein himself would accept this reformulation: he may, in the passage cited above, simply have been recommending us to look at how 'meaning' is used in ordinary language, in the same way that Austin was 'explaining the syntactics' and 'demonstrating the semantics' of 'meaning'. It is however very close to the view put forward in Ryle's *Ordinary Language,* 1953).

If meaning is language internal, what can one then say about the relation between language and the external world? The external world is clearly there: chairs can be sat on and windows broken. But it would seem to follow from this view that language imposes structure on the world and not vice versa, and that a particular language will

determine a particular structuring of the world.

An alternative way of approaching the question, that followed by Quine, will, in the end, lead to the same conclusion. The 'independent entities' view rather naturally leads to a pre-occupation with explanations of meaning in terms of referring, denoting, naming, etc. In the paper quoted at the beginning of this section, Quine examines the notion of referring and argues strongly that even in cases of ostensive definition – where the meaning of a word is explained by pointing to what the word refers to – it is in principle impossible to know precisely what is being pointed at. Thus, if I explain 'rabbit' to you by pointing to a rabbit, how can you know that I am pointing to the rabbit as an individual in the external world and not to, say the rabbit's ears, or to the rabbit as an example of a baby rabbit or a female rabbit. This is what he calls the 'inscrutability of reference'. From this he argues that reference is 'indeterminate'' in the sense that one cannot know what is being referred to, and that, a fortiori, translation too is indeterminate. In order to avoid the (absurd) conclusion that 'there is no difference between the rabbit and each of its parts or stages' he finishes by proposing that we

> begin by picturing us at home in our language, with all its predicates and auxiliary devices. This vocabulary includes 'rabbit', 'rabbit part', 'rabbit stage', 'formula', 'number', 'ox', 'cattle'; also the two-place predicates of identity and difference, and other logical particles. In these terms we can say in so many words that this is a formula and that a number, this is a rabbit and that a rabbit-part, this and that the same rabbit, and this and that different parts. In just those words. This network of terms and predicates and auxiliary devices is, in relativity jargon, our frame of reference, or co-ordinate system. Relative to it we can and do talk meaningfully and distinctively of rabbits and parts, numbers and formulas.

So once again, we are operating within a language, we are 'acquiescing in our mother tongue and taking its words at face value'. Both these alternative views, then, leave us with some uncomfortable questions about translation. If meaning is language internal, and if we operate meaningfully only within a frame of reference given by a particular language, what happens when we translate?

Quine hints at one possible way out when he talks of 'predicates of identity and difference, and other logical particles'. Perhaps there is a way out via logic and truth functions: through provision of a formal calculus within which we could model the external world, and such a model would be independent of any particular language. This is the line taken by, for example, Cresswell, in the excerpt below; where he tries to re-formulate the 'independent entities' view in terms of possible world semantics:

> Let us say that a and b (in different languages) are correct translations, each of the other, if they have the same meaning. For this definition to be viable we require language-independent entities to be the meanings of a and b, so that we can say that the entity which is a's meaning in its language is the same entity as that which is b's meaning in its language.
>
> Our theory of semantic competence enables a speaker to match up a and b with sets of possible worlds. Since possible worlds are language-independent he has the ability to tell whether a and b are logically equivalent, and it will be a necessary, though not in general a sufficient, condition

for a sentence a and a sentence b to be (correctly) intertranslatable that they be true in precisely the same set of possible worlds...

What seems likely is that an adequate treatment of these problems in possible-worlds semantics will require the use of theoretical entities which can represent distinct, though logically equivalent, propositions. If these entities are language-independent (as sets of possible worlds are language-independent) then, by treating them as the meanings of sentences in any language, we can say that a and b, in different languages, are correct translations of each other if they have the same meaning.

Since an adequate truth-conditional semantics will have to account for expressions like 'means the same as' it follows that, in order to solve its own internal problems, a theory of meaning based on the truth-conditional view of semantic competence offered in this paper will be sufficiently constrained to test the correctness of any synonymy claim in a single language; and granted the use of language-independent meanings, will therefore be able to test the correctness of any translation between one language and another.

Although this account avoids some (although not all) of the difficulties in the 'independent entities' view by offering a definition of them in terms of equivalence with possible worlds, it critically relates both meaning and translation to the formulation of truth conditions. This seems to me to be a mistaken view of what translation is, in that it neglects its essential linguistic character. Presumably, 'there is a chair in this room' and 'there is an article of furniture in this room known as a chair' have the same truth conditions, but it seems implausible to regard either one as a translation of the other. Even if this difficulty could be overcome, we are, in a rather similar way, left with the problem of formulating truth conditions for utterances containing semantically similar but translationally different terms in such a way that they can be distinguished. (For example, 'kill', 'slay', 'assassinate', 'murder'.) I find it difficult to imagine a formal apparatus which would allow us to do this in any very satisfactory way.

Thus, it seems that we are left with the notion that meaning is only describable in terms of a word's use within a language; there is no escape from words towards language independent entities. Yet we know that people do in fact translate, and that sometimes, at least, translation seems to be satisfactory. I want to suggest that the trick to understanding how this can be so is to take seriously the idea that a particular language imposes a structure on the world, rather than reflects a structure which is independently given; then, since the world as perceived through one language has much in common with the world as perceived through a different language, the structures imposed will sometimes, although not always, correspond in quite important ways. Learning a language does not involve taking one language and learning to express its terms and relations in terms of a second language, but involves learning a new world. Once this is done, it then becomes possible to identify correspondences between the two worlds created, as it were, by the two languages. On this thesis, learning a second language is not very different from learning a first language. Some support can be found for this in the empirical fact that adults learn languages with much less facility than do children: if language learning were simply a matter of learning a different,

but equivalent, vocabulary in which to express relations with which one was already familiar, then increased experience with one language gained through age should make it easier to learn the second language. Similarly, there is some (anecdotal rather than systematic) evidence which, in going against the language independent meanings thesis, tends to support the view put forward here: when bi- or multi-linguals attempt to recount a story or joke heard in one language in another, they very frequently hesitate to search for translations, suggesting that they remember the story in the language in which they heard it. (There is, however, some psychological evidence that humans do not seem to store surface language; see, for example, Johnson-Laird, 1974.) I do not know how much the question of mono-linguality vs. multi-linguality affects the issue or even whether the contradiction is, in fact, very deep).

A translator, then, identifies correspondences between the structures imposed by two languages; where such a correspondence exists, translation is possible, where it does not, only an approximation can be made. On reflection, this should not seem very shocking. There are very well-known cases where languages do not correspond; cases of lexical holes, for example, where one language simply does not have a vocabulary item corresponding to a vocabulary item in another language, or cases like colour words where there is strong evidence that different languages carve up the world in different ways. It does however have some consequences for the status of a variety of tools used within machine translation systems.

3   Tools for Semantic Analysis and the Search for Universals

In this section I want to look at a number of tools used within machine translation systems and ask what is their conceptual status in the light of the foregoing.

The clearest case is perhaps that of frames or scripts, when these are used as a way to drive the analysis rather than (or as well as) as a representation to be aimed at as the output of analysis. The classic case is to take some stereotypical situation or event, describe it in a formal calculus of some sort, and then carry out the analysis by seeking to relate elements of the input text to elements in the stereotypic description. The result will be a representation of the text, usually in terms of the stereotypic situation or event. In a machine translation system, the output translation is then based on this representation. (Metzing, 1979 is a collection of papers on the use of frames where the interested reader will find much more detailed description.) Despite some rather rash claims to the contrary (not the responsibility of the originator of the proposal), such organisations of knowledge about situations or events are very clearly bound to a particular culture and therefore to a particular language. Thus, translation can only succeed when there is a correspondence, in the sense of the preceding section, in the perceptions of the stereotyped event. Otherwise, the correspondence may be close enough to allow a frame for one language to be mapped onto a corresponding frame for the other. This would be the case, for example, with Minsky's birthday party frame if the system were translating between American English and French. (The typical party games are different, the food is different and so on, but the mapping could be done). In the extreme case, no correspondence will exist, and translation, as such, will be impossible. It is not too difficult, for example, to imagine a society in which children simply do not have birthday parties. Here, if a description of a birthday party occurred in a text to be translated, it

might be possible to add some explanation along the lines of 'In America, once a year on the date on which they were born, children are invited to a feast by their parents. This is called a birthday party.' But doing this is very clearly adding to the target language and its concepts, rather than taking it as given and simply moving from one language to another.

Equally clearly, the frames identified as aids for the analysis of a language can make no serious claims to universality in principle. It would perhaps be possible to imagine that some set of frames contingently applied to all languages (indeed, it was once, not very seriously, suggested that birth-life-death might constitute such a frame), but this would only be by accident rather than by definition within the theory.

Frame systems are rather similar in spirit to systems based on model theoretic semantics. Within such systems, typically, the independent existence of some state of affairs is assumed as the entity denoted by a semantic representation, and the semantic representation is isomorphic to the state of affairs. (Most logical grammars provide examples of this type of semantics.) However, the state of affairs modelled is often some feature of the world, rather than events or situations. A good example can be found in much of the recent work in artificial intelligence and in some machine translation systems on the modelling of time. An attempt is made to set up a model expressible in terms of a formal calculus, relatively independent of a specific language (I say relatively because most model-theoretic representations contain language specific words, like 'before' or 'after'; however they may contain parts such as their structure and parts of the vocabulary which are intended to be interlingual, and, more importantly, they receive an interpretation in terms of the semantic domain they model.) Analysis of the input text then involves extracting from it those elements relevant to the semantic domain in question and mapping the information thus retrieved onto the formal calculus of the model.

Here the question of universality is more difficult. The issue turns not so much on the existence or non-existence of independent meanings as on the existence or non-existence of something like the Kantian universal categories: aspects of the world or of thinking which no language can escape talking about, and where, even if different languages talk about these categories in different ways, it is possible to set up a model comprehending and subsuming their different ways. A strict Wittgensteinian approach would, I think, force one to give this notion up, and there is indeed some evidence that we should do so (Oatley, 1977). It is perhaps a superstitious dread of being set adrift in a formless world that makes us reluctant to do so. Two points however are clear. First, it is extremely unlikely that all of language can be dealt with in this way. To see this, one only has to think of the difficulty of accommodating discourse phenomena or pragmatic factors into a model, or of the intricacy (impossibility?) of dealing with intention. Second, even if a model of some category can be set up, if we admit that different languages may talk about the model in different ways, the existence of the model does not in itself guarantee translatability between the languages. Once again, the only cases in which translatability could be guaranteed would be those cases in which we had a correspondence between the two languages in the sense of the first section.

Quite direct claims for universality have been made by the proponents of conceptual dependency theory: 'Conceptual Dependency theory is intended to be an interlingual meaning representation. Because it is intended to be language free, it is necessary in our

representation to break down sentences into the elements that make them up. In order to do this it is necessary to establish a syntax of possible conceptual relationships and a set of conceptual categories that these relate' (Schank & Rieger, 1974).
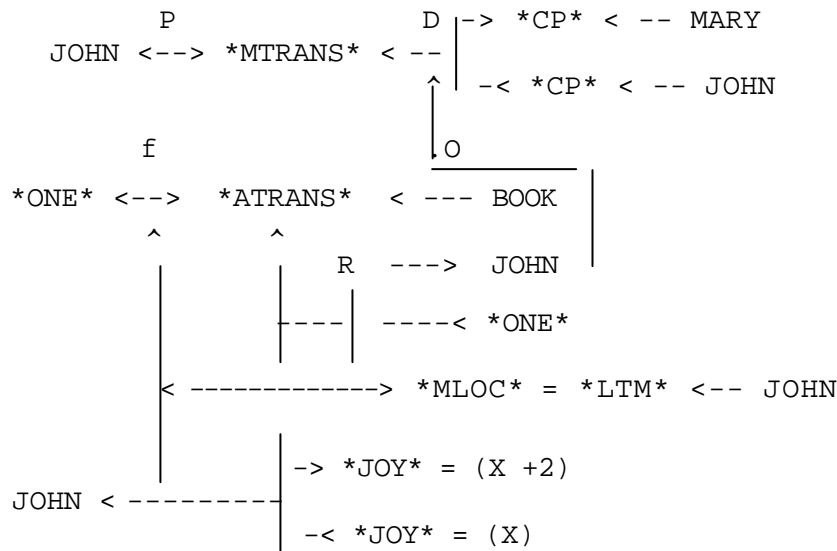
The primary conceptual categories are taken to be ACTs, real world actions, of which a small number are claimed to be primitive. (Twelve in the paper taken here as a basic reference: the number varies slightly in other papers.) Other conceptual categories, for example 'real world objects', 'attribute of actions', 'times', 'locations' are postulated, but are not worked out in any great detail. The theory concentrates on the set of primitive actions. With each action is (obligatorily) associated a sub-set of the conceptual cases (objective, recipient, directive, instrumental, where instrumental is itself a complete propositional structure) and a set of inferences. The inferences occur automatically, but are not guaranteed to be correct. Thus, within conceptual dependency theory, the meaning of 'John told Mary that he wants a book' will not only include John's transmitting to Mary that he would be happier if someone gave him a book, but will also include that John will also be happier if he can transmit information from the book to himself: i.e. the inference is made that John wants a book because he wants to read it.

The representation theory of conceptual dependency and the inferencing techniques associated with it can be considered independently, as can be seen from Rieger's thesis, which is mainly concerned with distinguishing different types of inference (Rieger, 1974). Here we shall consider only the status of the primitive acts and of the 'conceptualizations' (i.e. propositional structures) of which they form a part.

First, and most obviously, primitive acts themselves cannot be considered to be independent meanings of words: the mere fact that (at least) all the verbs of English can, it is claimed, be mapped onto twelve or so primitive acts makes this obvious.

Nor would Schank argue differently. In Schank (1975) he postulates the existence of two kinds of memory. (Conceptual dependency theory was originally developed as a model of human memory: its use inside natural language understanding systems was seen as a way of testing the theory, rather than as an independent enterprise. Later workers, however, have sometimes taken over the theory as a theory about language processing.) One of these is conceptual memory, which is structured in the conceptualizations described briefly above. The other is a lexical memory, which is said to contain 'all of the information about words, idioms, common expressions etc.' and which 'links these to nodes in a conceptual memory, which is language free'. Furthermore, the lexical memory seems to have some structure, since, in discussing the notion of 'superset' in memory (in relation to semantic network representations such as those of Collins & Quillian, 1969, where 'bird', for example, is represented as a superset of 'canary', 'ostrich' and so on), Schank claims that the number of such supersets is very small and that they are 'mostly artificial constructs with definitions in lexical memory' It is not clear, though, where or how this lexical memory comes into being: Schank (1974) seems to contrast it with conceptual memory; 'Once we change semantic memory by separating out lexical memory, we are left with a set of associations and other relations between concepts that could only have been acquired by personal experience'. This would seem to imply that lexical memory is not acquired by personal experience, but nothing more is said, and it would perhaps be rash to push too far on the basis of a single sentence.

On the other hand, the conceptualizations themselves claimed to represent the meaning of sentences, and thus could be expected to explicate or at least subsume the meaning

47

```
              P                 D  -> *CP* < -- MARY
    JOHN <--> *MTRANS*  < --|
                           ^ |  -< *CP* < -- JOHN
                           | |
              f            | O
                           |_____
  *ONE* <-->  *ATRANS*  < --- BOOK       |
         ^       ^                        |
         |       |    R  --->  JOHN       |
         |       |---|                    |
         |       |   | ----< *ONE*        |
         |       |   |                    |
        |< -------------> *MLOC* = *LTM* <-- JOHN
         |
         |      |-> *JOY* = (X +2)
  JOHN < -------|
                |-< *JOY* = (X)
```

(i.e., the conceptualization to the left has
      mental location John's LTM)

of the individual words. That they do not in fact do so in any normal sense of 'meaning'
can be seen by considering the conceptualization for 'John told Mary that he wants a
book' (from Schank & Rieger, 1974).

There is no need to go into detail. The essential point is that two primitive acts,
ATRANS and MTRANS are involved, the first involving 'the transfer of an abstract
relationship such as possession, ownership or control', the second 'the transfer of mental
information between animals or within an animal'. It is difficult to see how, with the
limited apparatus offered, it would be possible to distinguish between different ways of
telling Mary, or of being given a book. If John wrote to Mary, one could perhaps capture
this by including an instrument case: but if he hinted it, or informed her that..., or insisted
to her that..., how could this be represented? Thus, conceptual representations of this
sort actually carry less meaning than the sentences they represent. Schank, however,
does not base his argument on meaning but on 'information': 'Information is not lost by
the use or primitive ACT's' (Schank, 1975).

A defendant of the theory might argue therefore that we were being finicky: that
identifying and preserving all important information was all that mattered. The claim
then would be that it was possible to set up a formal model which captured all information
necessary to communicate what we might call kernel meaning - the bare bones without
any of the niceties of nuance or of subtlety. Whether or not such a model were meaning-
preserving would then, of course, be very contentious, and would depend critically
on a definition of what the essential information was; which, in its turn - and this is
logically prior - depends on a belief that it is, in some way, possible to identify, language
independently, what has to be communicated.

48

Schank himself seems to believe this to be possible, although he does not claim that there is some fixed number of ACT's which will constitute the 'right' set:

> There is no right number of ACTs. It would be possible to map all of language into combinations of mental and physical MOVE. This would, however, be extremely cumbersome to deal with in a computer system. A larger set (several hundred) would overlap tremendously causing problems in paraphrase recognition and inference organization. The set we have chosen is small enough not to cause these problems without being too small. Other sets of the same order of magnitude might do just as well.

All of this, it seems to me, is to make the explanation of meaning more problematic rather than less. Now it is not even possible to look at how a word is used within a language: instead we have to search for some common element in the use of a group of words (perhaps deciding rather arbitrarily what the members of that group are), define that common element and embed it into an information structure. Having done so, we can reasonably ask what that information structure means. Only two answers seem possible. The first is an answer in terms of some set of operations performed on that structure by, say a computer program, in which case the adequacy of the structure is, presumably, to be judged by the output of the program. To do this, some metric must be set up in terms of which the judgment can be made - and this begins to sound suspiciously like a vicious circle, in that we are back to asking whether the program mimics use in ordinary language. The second is an answer already in terms of ordinary language, without passing through explanation in terms of the computer program, when the circle becomes evident.

Since conceptual dependency was originally developed as a memory model, another justification could be sought in the psychological reality of the conceptualizations and their organization. This too seems extraordinarily difficult to prove, since it involves comparing two unlike objects. The structure of the brain and its workings is still very largely a mystery. The designers of experiments designed to show psychological reality are therefore driven back on an experimental design whereby some particular input is claimed to predict some particular output. Unfortunately, even when the prediction is correct, such an experiment can say little, if anything, about how the output was obtained. An abacus and a computer can be claimed to take similar inputs and produce similar outputs, but no-one claims that they do it the same way.

As far as translatability goes, it should be clear that conceptual dependency theories simply cannot be used to translate. They are, by their nature, reductionist, in that they lose much of what is normally considered meaning: even if we take cases where not much is lost, 'Jean a acheté une voiture' cannot be translated as 'Someone sold John a car'. In other words, at best such theories lead to paraphrase systems. At worst they lead to re-expressing the full vocabulary of French, say, in the eight hundred or so words of basic English.

Associated with the primitive acts of conceptual dependency theory are, as we have already seen, four conceptual cases. The use of deep cases had become quite common in language processing systems since the appearance of Fillmore's seminal paper in 1968. (Although it seems that Japanese linguists – perhaps not surprisingly given the structure of Japanese – were working along similar lines well before Fillmore). Fillmore makes

a distinction between surface case, marked in some languages by inflection (e.g. nominative, accusative, genitive, dative, ablative in Latin), by prepositions in others (e.g. English), by post-position particles in others (e.g. Japanese), and 'deep' case, which is held to capture the semantic role of the participant in a predicate. In terms of deep case 'John', 'the door' and 'the key' preserve their semantic role in all of the following sentences, despite the differences in surface structure:

'John opened the door with the key'
'The key opened the door'
'The door opened with a key'
'The door opened'
'The door was opened by John'
'The door was opened by John with a key' etc.

Both for Fillmore and for Schank, the number of deep cases is fairly small. Workers in machine translation have tended, on the other hand, to work with much larger case sets (typically of around twenty to thirty cases: the increase in the number of cases comes largely from finer classification of circumstantials to take in roles like 'result', 'cause', 'concessive' in addition to the time and location cases).

Fillmore himself claimed only that the notion of deep case was universal, in the sense that, in Chomskyan terms (Chomsky, 1965) a notion of a case formed part of the base structure: '...what is needed in a conception of base structure in which case relationships are primitive terms of the theory and in which such concepts as "subject" and "direct object" are missing'. However, he makes no claim for the universality of particular case sets: 'My claim is, then, that a designated set of case categories is provided for every language, with more or less specific syntactic, lexical, and semantic consequences, and that the attempt to restrict the notion of "case" to the surface structure must fail'.

(Fillmore did, though, rather hope that on investigation, some universality might emerge: 'It seems to me that if there are recognizable intrasentence relationships of the types discussed in studies of case system... that if these same relationships can be shown to be comparable across languages and that if there is some predictive or explanatory use to which assumptions concerning the universality of these relations can be put, then surely there can be no meaningful objection to using the word case, in a clearly understood deep-structure sense, to identify these relationships'. We shall return to the question of empirical investigation in the next section).

Some support for the idea that different languages rely on different case sets can be found also in the empirical fact that the Japanese national machine translation project, which relies heavily on the use of deep case, uses a case set for the analysis of Japanese which differs slightly from that used in the same project for the analysis of English (Nagao et al. 1985). Even where the deep case intuitively seems almost to correspond to one of the universal categories (as we called them earlier), there is evidence that different languages give a somewhat different semantic content to the case roles. Consider, for example, space relations in English and French, where the choice of preposition is determined by quite different factors:

in (cities, countries, enclosed spaces)
on (islands, mountains, streets)
at (buildings)

vs   à Paris   en Avignon (phonetic)
en France   au Japon (gender)
en Corse   sur l'île d'Elbe (political status)

(This example is due to H. Somers).

In view of what was said in the first section, it should not greatly surprise us to find doubts about the a priori universal status of deep cases or semantic relations. They could only be universal if all languages structured the world in the same way, which seems unlikely if language is taken as imposing structure on the world rather than as reflecting a world already structured. As with other aspects of meaning, sometimes there will be a strict correspondence in the relationships perceived, sometimes not. Translation will consist, as before, in identifying and using the strict correspondences, learning to massage the near correspondences and finding, where possible, other ways of expressing the mis-matches.

One final tool of semantic analysis should be considered before we leave this topic: the use of semantic markers. Semantic markers can be thought of in two quite distinct ways, although there is a frequent and unfortunate confusion between the two. Both can be found in the original Katz and Fodor proposal. First, semantic markers are intended to represent the meaning of individual words. Second - and, conceptually, this is quite a distinct notion - semantic markers are used to encode selectional restrictions, and thus to inhibit the co-occurrence of certain words. As an example, take the dictionary entry for "ball"

Ball → concrete noun → (social activity) → (large) → (assembly) → [dance]
Ball → concrete noun → (physical object) → [sphere]
Ball → concrete noun → (physical object) → [cannon-ball]

(From Katz and Fodor, 1963.)

The objects in round brackets are the semantic markers, the objects in square brackets 'distinguishes'. The basic notion of selectional restrictions is very simple: it rests on the observation that some word senses which are fundamentally predicates or relations may only predicate on or relate objects having certain properties. Thus, the verb 'kick' for example, in the ordinary sense of propel by use of a foot or a hoof requires an (animal) as subject and a (physical object) as object. So, in the sentence 'The page kicks the ball', the sense of page as the page of a book is ruled out by the requirement that the subject must have a marker (animal), and the first of the three senses of ball in the example is ruled out by the requirement that the object must have (physical object).

The distinguishers are intended to contain the remnant of the meaning of a particular word sense which is not already accounted for by the semantic markers in the entry for that sense. They do not, though, interact with the selectional restrictions, which operate only in terms of the markers. However, markers and distinguishers together should capture the meaning of the word.

51

If we now neglect the distinction between markers and distinguishers, and consider only the two mechanisms, that of defining the meaning of the word, and that of serving as clues for disambiguation and thus inhibiting certain co-occurrences, we see immediately that the two functions are quite different.

Taken as a means of defining the meaning of a word, markers become semantic primitives, combinations of which produce particular meanings. This is a variation on the 'independent entities' theory of meaning, where some set of primitives is taken as primary (and independent) and 'meanings' constructed out of them. As such, it is open to all the objections levelled against that theory.

Wilks (1975a & 1975b) has put forward a version of markers-as-meanings which avoids this objection. There he argues that semantic primitives, in such a scheme, have exactly the same status as words of the language, and that, in fact, they coincide with words of the language in everything except that they are words accorded an extra privilege in that verbal explanation of the meaning of other words is only allowed in terms of the privileged words.

If we take this position, it follows that there can be no 'right' set of semantic primitives. To quote Wilks (1975b): 'It follows from this that there can be a variety of primitive languages for semantic descriptions, no one necessarily better or worse than any other, any more than my vocabulary is better or worse than yours if I know 100 English words you don't, and you know 101 that I don't. In the case of each primitive vocabulary, the only ultimate test will be the success or failure of linguistic computations that make use of it'. He goes on to remark that there are some limits to the parallel between a primitive vocabulary and ordinary language, for example that a primitive vocabulary should not have synonymous primitives, and earlier in the paper suggests other 'operational' criteria for the selection of a primitive vocabulary.

There are further objections to the notion of defining a single, universal, correct set of primitives. The first is already implicit in the quotation from Wilks above. If a set of primitives is to be universal, each primitive must be interpreted in the same way. But how is this to be achieved? It seems difficult to offer a definition, since primitives are primitive by definition, and are those terms in a theory which are no longer defined. Therefore, all that could be offered would be an explanation of the use of the primitive, and if this is offered in words or by examples first no consistency of interpretation can be guaranteed and – worse – we are again in the vicious circle where primitives are defined by words and words by primitives. Secondly, if we accept that there can be no objective correlative of a would-be semantic primitive, how could anyone using a set of primitives ever know that he was using it in the correct way? In the end, the interpretation of a set of primitives can only be given by the use to which they are put within some formal system. This point too we shall return to in the next section.

The arguments offered here against the possibility of defining a correct set of primitives hold equally strongly for the definition of a set of markers to be used as clues in disambiguation if it is assumed that for every language, or even within one language, there will be a single correct set. This is sometimes somewhat obscured by the fact that the markers most commonly found in systems using them in this way tend to constitute a rather small set using words like 'animate', 'physical object' which reflect common strong correspondences between languages. As soon as the set is enlarged, the temptation to believe in a single universal set recedes.

One final remark: if we accept Wilks' thesis that primitives are no more than ordinary words accorded a privileged status, then of course they cannot be universal across languages. 'Animate' is an English word, and enters into the nexus of English language use; 'animé' is a French word entering into the nexus of French language use. The two uses do not correspond.

4 The Proper Place of Semantics in Machine Translation

The reader might be forgiven for believing that I am advocating the abandonment of semantic tools for use in machine translation systems. In fact, this is not so, I am advocating only that they are seen in a correct perspective, and that system designers do not prejudice their own intellectual enterprise by trying to insist on a universality which does not exist. In this section, I want to put forward some suggestions about the appropriate use of semantic tools. Apart from some very brief remarks on the other techniques discussed in the last section, I shall concentrate mainly on the use of case and of semantic features.

Let me start by insisting once again on the notion of translation as a linguistic undertaking, where the primary task is to find words of one language, embedded in an appropriate syntactic structure, which capture as closely as possible the correspondence between the world of the target language and the world of the source language, this latter itself described via the mechanisms of the source language, with all its richness of vocabulary and its subtlety of structure.

Providing that this is not forgotten, with the result that we finish up with an impoverished paraphrase rather than a translation, there is an obvious place for the use of frames whenever there is a very strong correspondence between the target language world and the world of the source language. There are, of course, other technical problems in the use of frame-driven systems which have not been discussed in this paper which would then become relevant, but they do not matter for the point being made here.

The case of conceptual dependency type theories is rather more difficult. In the preceding section, I have suggested that they are based on a reductionist hypothesis of meaning, in that whole classes of verbs are mapped onto the primitive acts, with a consequent automatic loss of meaning. This would seem to make them unsuitable for use in a translation system if the conceptual dependency representation is taken as an interlingua (of sorts) through which translations must pass. (Recent work using conceptual dependency within the Yale group has tended to concentrate on frame-like systems, using conceptual dependency representations as the smallest building blocks of the frames. This may make it seem that this paragraph is inconsistent with the last. But, there is, of course, no essential intimate connection between frame systems and conceptual dependency representations, as witness the fact that most frame systems use quite different, and quite varied, representations.)

On the other hand, although we have not discussed the issue much in this paper, conceptual dependency representations are often used as the basis for inference making. Perhaps, then, they could be useful in those cases where inference becomes important, for example to resolve difficult problems of pronoun reference. But I would want to claim that their use should be restricted to the solution of such problems, and that they

53

should not be seen as the primary tool around which a system should be built, thus replacing the nuts and bolts of standard linguistic analysis.

Model theoretic semantics too may find a place for sub-parts of a system where it is possible to set up a formal calculus onto which some aspects of the language(s) treated by the system can be mapped, but cannot serve as the basis of a system in general.

The difference between what I am suggesting here and the proposals normally found in the literature is this: normally, the use of frames, of conceptual dependency like theories or of model theoretic semantics is seen as an alternative to a conventional linguistic analysis based on syntax. Here I am suggesting that such tools should work in co-operation with a more conventional model. By this, I do not want to suggest that the only possible semantics is an interpretative semantics, whereby semantic representations are considered to be straightforward projections of syntactic representations. Typically, within such theories, the semantic representation is derived from autonomously defined syntactic structures, there are quite severe constraints on the rules which define the mapping from one to the other, and in consequence, the semantic representations tend to be quite similar to the syntactic representations. (For examples of this type of theory, see Chomsky, 1981, Jackendoff, 1983, Bresnan, 1982.) Since, within a machine translation system, I take it to be the main task of analysis to neutralize the differences between languages as much as possible, and since, within such theories, the semantic interpretation tends to reflect the syntactic analysis, which, in its turn, reflects differences which could, otherwise, be neutralized, I take interpretive semantics to be of only limited interest within a machine translation system. Rather, I am suggesting some kind of 'blackboard' approach, where different kinds of analysis contribute their conclusions to a common pool, and the final result is constructed on the basis of what is found in the common pool.

The two remaining semantic tools of those discussed in the last section have normally been perceived as additions to, or aids towards, a conventional linguistic analysis, rather than as a replacement for it. Most commonly, deep cases are seen as a way to neutralize the contingent syntactic variations of surface structure: a constituent analysis of some sort is performed, and the constituents thus identified mapped into a set of case roles, which are intended to be common to the two or more languages being treated within the particular machine translation system. This indeed seems to be the most appropriate use for a set of deep cases. The difficulty, of course, comes with the definition of an appropriate set, especially if we accept the argument of the last section that no correct universal set can be defined a priori. To regard this as an insurmountable obstacle seems to me, however, to be unnecessarily defeatist, since it is perfectly possible to define an operational procedure whereby correspondences between cases can be found across languages. Such a procedure would consist in establishing a set of cases for each language separately, taking as criteria for an appropriate set such factors as results of previous work on that language, the possibility of ensuring consistent assignment (by men and by machines), the distinctions made within that language, and so on. Transfer would then involve quite explicitly mapping between the deep case roles of one language and those of another. If it were found, on the basis of such an empirical investigation, that the mapping between any two case roles was always one to one (or even, with a very slight adjustment of the case set, could be made so), it could then be taken that a relationship common to both languages had been found, and explicit transfer could be

54

dropped.

(As an historical, and perhaps provocative note, I believe Fillmore himself to have been assuming the necessity of just such an empirical investigation, as shown by the quotation in the last section and by sentences such as 'I am going to suggest below that there are many semantically relevant syntactic relationships involving nouns and the structures that contain them, that these relationships ... are in large part covert but are nevertheless empirically discoverable, that they form a specific finite set, and that observations made about them will turn out to have considerable cross-linguistic validity. I shall refer to these as "case relationships"'.) One of the beauties of a transfer-based machine translation system is that it provides a clear framework in which such an empirical study can be carried out.

It is not nearly so clear that any similar empirical study can be proposed in the case of semantic features. It will, perhaps help to clarify the issue somewhat if we distinguish two possible places in which semantic features might be used. The first would be their use inside some area of language susceptible to modelling in the model theoretic sense described earlier (if such an area exists). If we take as an hypothesis that it is possible to set up such a model, for time say, then it is easy to imagine that part of the model itself would consist in a set of features. In such a case, there is no question of an appropriate or inappropriate set: the model itself determines the features which are part of it. But there is another common use of semantic features which consists in using the features as an aid towards some further result, in carrying out the mapping between the text and the model, say, or as an aid to lexical or structural disambiguation. This case, I would argue, is quite different. Here, the semantic features can be seen most fruitfully as an attempt at semantic categorisation, and thus closely parallel to syntactic or morphological category names. If this is so, then just as syntactic categories receive their interpretation within the linguistic theory which makes use of them – in machine translation terms via the grammar rules in which they are used – so do semantic categories. And just as we expect that different theories will make use of different syntactic categories, and do not ask for there to be one universal set, so we should expect that different theories will make use of different semantic categories, and not ask for or expect universality here either. We would then take Wilks' proposal, whereby semantic primitives are regarded as a primitive vocabulary, as a proposal for establishing a preliminary set of semantic features within a particular theory as well as within a particular language, and would add to his criteria of lack of redundancy and disjointness the further criterion of usefulness within the theory. The semantic categories would then serve as aids to establishing a correct analysis and a correct generation of a specific language, and would serve cross-linguistically only as extra information towards making correct choices in transfer.

5   Conclusion

This paper, starting from philosophical considerations, has discussed the status of a variety of tools used in natural language understanding systems in general and in machine translation systems in particular, with the aim of determining their most appropriate use within such systems. Relatively little has been said that will be new to philosophers or linguists: the point in saying it has been to emphasize to workers in machine translation and in computational linguistics in general the importance of being clear about the nature

of the intellectual tools they use and about their reasons for using them. That point will remain valid even if every argument or conclusion contained in the present paper proves to be mistaken.

Acknowledgments

References

[1] Austin, J.L., 1961. "The Meaning of a Word," *Philosophical Papers.* Oxford University Press.

[2] Austin, J.L., 1962. *How to do Things with Words.* Harvard University Press, Cambridge, MA.

[3] Bresnan, J. (ed.), 1982. *The Mental Representation of Grammatical Relations.* MIT Press, Cambridge, MA.

[4] Caton, C.E., 1971. "Philosophy: Overview," in Steinberg & Jakobovits, 1971.

[5] Chomsky, N., 1965. *Aspects of the Theory of Syntax.* MIT Press, Cambridge, MA.

[6] Chomsky, N., 1981. *Lectures on Government and Binding.* Foris Publications, Dordrecht.

[7] Colling, A.M. and Quillian, M.R., 1969. "Retrieval Time for Semantic Memory," in *Journal of Verbal Learning and Verbal Behavior*, 8.

[8] Cresswell, M.J., 1978. "Semantic Competence," in Guenthner and Guenthner-Reuter, 1978.

[9] Fillmore, C.J., 1968. "The Case for Case," in Bach and Harms (eds.), *Universals in Linguistic Theory,* Holt, Rinehart and Winston, New York.

[10] Guenthner, F. and Guenthner-Reuter (eds.), 1978. *Meaning and Translation.* Duckworth.

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

[12] Johnson-Laird, P., 1974. "Memory for Words," *Nature.*

[13] Katz, J. and Fodor, J.A., 1963. "The Structure of a Semantic Theory," *Language,* 39.

[14] Metzing, D. (ed.), 1979. *Frame Conceptions and Text Understanding.* De Gruyter, Berlin.

[15] Minsky, M., 1981. "A Framework for Representing Knowledge," in Haugeland (ed.), *Mind Design.* MIT Press, Cambridge, MA.

[16] Nagao, M., Tsujii, J.-I. and Nakamura, J.-I., 1985. "The Japanese Government Project for Machine Translation," *Computational Linguistics,* 11(2-3).

[17] Oatley, K.G., 1977. "Inference, Navigation and Cognitive Maps," in Johnson-Laird and Wason (eds.), *Thinking.* Cambridge University Press.

[18] Quine, W.V., 1960. *Word and Object.* MIT Press, Cambridge, MA.

[19] Quine, W.V., 1971. "The Inscrutability of Reference," in Steinberg and Jakobovits, 1971.

[20] Rieger, C, 1974. *Conceptual Memory.* Unpublished Ph.D. Thesis, Stanford University.

[21] Ryle, G., 1953. "Ordinary Language," *Philosophical Review,* LXII.

[22] Ryle, G., 1957. "The Theory of Meaning," in C.A. Mace (ed.), *British Philosophy in the Mid-Century.*

[23] Schank, R.C., 1974. *Is There a Semantic Memory.* ISSCO Working Paper No. 3.

[24] Schank, R.C., 1975. "The Theory of Meaning," in C.A. Mace (ed.), *British Philosophy in the Mid-Century.*

[25] Schank, R.C. and Rieger, C.J., 1974. "Inference and the Computer Understanding of Natural Language," *Artificial Intelligence* 5(4).

[26] Somers, H.L. (Forthcoming). *Valency and Case in Computational Linguistics.* Edinburgh University Press.

[27] Steinberg, D.D. and Jakobovits, L.A., (eds.) 1971. *Semantics.* Cambridge University Press.

[28] Wilks, Y, 1975a. "Primitives and Words," in Schank and Nash-Webber (eds.) *Proceedings of the Workshop on Theoretical Issues in Natural Language Processing.* Bolt, Beranek & Newman, Cambridge, MA.

[29] Wilks, Y., 1975b. *Seven Theses on Artificial Intelligence and Natural Language.* ISSCO Working Paper No. 17.

[30] Wittgenstein, L., 1953. *Philosophical Investigations.* Blackwell, London.

[31] Wittgenstein, L., 1958. *The Blue and Brown Books.* Blackwell, London.