

DEGREES OF UNDERSTANDING

Eva Hajičová and Petr Sgall

Faculty of Mathematics and Physics

Charles University

Malostranské n. 25

Prague 1, Czechoslovakia

1. Introduction.

Along with "static" or "declarative" descriptions of language system, models of language use (the regularities of communicative competence) are constructed. One of the outstanding aspects of this transfer of attention consists in the efforts devoted to automatic comprehension of natural language which, since Winograd's SHRDLU, are presented in many different contexts. One speaks about understanding, or comprehension, although it may be noticed that the term is used in different, and often rather unclear, meanings. In machine translation systems, as the late B. Vauquois pointed out (see now Vauquois and Boitet, 1985), a flexible system combining different levels of automatic analysis is necessary (i.e. the transfer component should be able to operate at different levels). The human factor cannot be completely dispensed of; it seems inevitable to include post-edition, or such a division of labour as that known from the system METEO. Not only the semantico-pragmatic items present in the source language structure should be reflected, but also certain aspects of factual knowledge (see Slocum, 1985, p.16). It was pointed out by Kirschner (1982, p.18) that, to a certain degree, this requirement can be met by means of a system of semantic features. For NL comprehension systems the automatic formulation of a partial image of the world often belongs to the core of the system; such a task certainly goes far beyond pure linguistic analysis and description.

Winograd (1976) claims that a linguistic description should handle "the entire complex of the goals of the speaker" (p.269,275). It is then possible to ask what are the main

features relevant for the patterning of this complex and what are the relationships between understanding all the goals of the speaker and having internalized the system of a natural language. It seems to be worth while to reexamine the different kinds and degrees of understanding.

2. Understanding the sentence.

Segmentation, disambiguation and identification of units of the individual levels are the main tasks of the elementary steps of understanding an utterance.

(i) The lowest step consists in the segmentation of the continuous flow of sound into individual phones; their sequence can be understood as consisting in subsequent points of a feature space, the individual feature values of the space corresponding to the distinctive features, which have to be identified. Disturbances on this level may be due to noise or to physiological irregularities. On the phone, in a crowded room, the utterance *I don't understand you* may mean that the hearer is unable to identify the uttered phones.

(ii) - (iii) A phoneme may consist of several phonic variants, and a string of phonemes can be decomposed into morphs, each of which corresponds to a morpheme; the latter is a feature space again, the values of the features here being the semes (preterite, genitive, plural, ...). Thus, if F (a sequence of phones) is (the phonetic shape of) an utterance, *Phone*, *Phoneme*, *Morph* and *Morpheme* being the sets of all phones, phonemes, morphs and morphemes, respectively, of the language described, we can write:¹

$$F = (f_1, \dots, f_n), \text{ where } 1 \leq n,$$

P is a mapping of *Phone* onto *Phoneme*
 $Morph \subset Phoneme^*$ (i.e. *Morph* is a proper
subset of the set of all strings
of phonemes)

$M \subset Morph \times Morpheme$

$Phoneme^0 = \{x \in Phoneme^*; \\ \exists k, \exists m_1, \dots, m_k \in Morph(x = m_1, \dots, m_k)\};$

thus $Phoneme^0$ is the set of strings of phonemes that constitute strings of morphs. The disambiguation identifying the string of morphemes conveyed by x can only be made, in the general case, after the syntactic patterning of the sentence, its meaning and its fitting into the co-text and situation has been grasped. The steps of understanding thus cannot be performed in a uniform order; they are checked by means of trial and error.²

(iv) If one is reading without paying much attention, one "wakes up" when one's more or less subconscious interpretation encounters an obstacle (e.g. with a garden-path sentence); one realizes that it is necessary to go back in the text to where one's attention was derailed, and read again, paying due respect not only to (surface) syntax, but also to understanding on the higher degrees.

(v) If the hearer understands the linguistic (literal) meaning (or, reaches a disambiguation of the utterance), s/he understands e.g. *this letter* as the Objective of (1) and as the Actor of (2); further detours (using criteria from higher degrees) decide on the role of *planes* in a token of (3).

- (1) This letter I got only today.
- (2) This letter came only today.
- (3) Flying planes can be dangerous.

The level of linguistic meaning (tectogrammatrics, underlying structure) is language specific and comprises the theta roles (deep cases, valency slots) as well as the topic/focus articulation (which is semantically relevant for the scopes of operators and for presuppositions); see Sgall et al. (in press).

Without knowing the situation it is impossible to tell who is referred to by I in (1), what is meant by *this letter*, and so on. Thus, considering the sense of an utterance (i.e., of a token of sentence in a discourse) to consist in a combination of the meaning of

the sentence with the specification of the reference of the expressions it contains, we come to a further degree of understanding, illustrated by *Are you speaking about the letter you got from my brother?*. This step leads us beyond the system of language, which has no means to identify the objects referred to. In the prototypical situations of communication I is understood, since who hears the utterance knows who utters it. *You, here, now, we* (and thus also *your, up to now, ...*) are similar, although they are not delimited as for the range of reference. Without knowing the situation, the hearer also is unable to specify the reference of *this letter, the house, a friend of mine...* The sense of utterances can be identified only by means of non-linguistic clues.

3. Understanding in communication.

(vi) The identification of reference is conditioned by non-linguistic factors, with all expressions not having a unique reference. The main factor is the speaker's assumption concerning the hierarchy of salience (prominence) of the items in the hearer's memory. As Hajičová et al. (1982; 1984) point out, *it* refers to the most salient item, *the table* to the table activated by having occurred in the focus (comment) of a preceding utterance (or by situation, common interest,...)

(vii) The next degree concerns habitual connotations, a possibly intended inference, see Winograd (1976, 275), regular cognitive relationships (cf. 'frames' and 'scenarios'), and issues connected with conversational maxims and stone walling, see Joshi et al. (1984)

(viii) Non-habitual inferences are placed along a scale of complexity, cf. Hintikka (1975), so that an elementary use of intellect (proper to most human beings, though not *qua* users of a language) may be distinguished from conscious intellectual effort.³

Another hierarchy of inferences, concerning the difference between 'what I am told' and 'why', starts with the distinction of direct and indirect speech acts, and continues with that between illocution and perlocution, including further degrees of the type "He

wants me to react in this way; but why does he?" - "He wants to achieve this and this by my reaction; but for what purpose?",...

4. Conclusions.

The theory of language cannot be exclusively based on language understanding. Coming back to the question put in § 1, we find that among the degrees of understanding only those from § 2 immediately concern the structure of language, and even with them factual knowledge plays a big role. The degrees (vi) to (viii); and thus also "the entire complex of goals" of the speaker goes far beyond the domain of linguistics, contrary to Winograd. A theoretical account of language is a necessary ingredient of a model of comprehension; it allows us not to use ad hoc solutions, which at a later stage could prevent a useful generalization of the comprehension system, its adaptation to new applications, etc. When evaluating a linguistic theory one should ask whether it can be embedded in a theory of communication; an economical account of topic and focus makes it possible to describe the meaning of a sentence as a procedure instructing the hearer how to change the contents of her/his memory, and thus to connect the handling of sentence structure with that of the patterning of a discourse.

Notes

- 1 We neglect the cases where a phone functions in different contexts as a variant of two different phonemes.
- 2 Disambiguation was discussed in the frame of neural-net linguistics and cognitive science by Schnelle (1984, esp.12); cf. his ex. *Der Patient hatte einen Wachtraum vs. Die Kaserne...*
- 3 Other aspects of inferencing are studied as based on logical entailment, leading from the sense of an utterance - cf. 2 (v) - to the proposition (a function from possible worlds into truth values); the specification of reference mostly is tacitly assumed to be present in a proposition. For an analysis of belief sentences and other "propositional"

attitudes, as well as of such paradoxes as that of the Liar's and for contradictions such as those concerning round squares and similar expressions it is indispensable to work with a clear difference between (a) the level of linguistic meaning (disambiguated underlying structure), (b) the layer of sense (including the specification of reference), and (c) the psychological domain of context (requiring a description of the relative salience of the items in the speaker's and hearer's memories); cf. Sgall et al. (in press, Ch.1).

References

- Hajičová E., Sgall P. and J.Vrbová (1984), Topic, Focus and How to Identify Them, Groningen Arbeiten zur germanistischen Linguistik 25, 121-143
- Hajičová E. and J.Vrbová (1982), On the Role of the Hierarchy of Activation, in COLING 82 (ed.J.Horecký), Prague, 107-113
- Hintikka J. (1975), Impossible Possible Worlds Vindicated, Journal of Phil. Logic 4, 475-484
- Joshi A., B.Webber and R.M.Weischedel (1984), Preventing False Inferences, in Proceedings of Coling 84, Stanford, 134-138
- Kirschner Z. (1982), A Dependency-Based Analysis of English for the Purpose of Machine Translation, Explizite Beschreibung der Sprache und automatische Textbearbeitung IX, Prague
- Schnelle H. (1984), Programming of Net Linguistic Processes, GENET 17a, Bochum
- Sgall P., Hajičová E. and J.Panevová (in press), The Meaning of the Sentence in Its Semantic and Pragmatic Aspects, Reidel (Dordrecht) and Academia (Prague)
- Slocum J. (1985), A Survey of Machine Translation, Computational Linguistics 11, 1-17
- Vauquois B. and C.Boitet (1985), Automated Translation at Grenoble University, Computational Linguistics 11, 28-36
- Winograd T. (1976), Towards a Procedural Understanding of Semantics, Revue internationale de philosophie, 260-303