Semantics as lexicography

Bart Geurts

University of Nijmegen

Over the past few decades, the interpretation of language has gradually come to be accepted as a subject deserving its own academic regalia, such as conferences, journals, and university chairs. The definition of the emerging field of inquiry is, as yet, very moot. It is not even decided that it is a single field, rather than several related ones, and whether it (or they) should be subsumed under linguistics, psychology, or somewhere else. But there is a measure of consensus about the kind of topics the fledgling discipline is supposed to address. Here are a few examples:

Quantification
Quantified expressions like 'all', 'most', and 'many' raise a host of interdependent questions concerning interpretation, their scope-taking properties, the way they interact with anaphoric expressions, and so on. Quantification is without any doubt one of the major topics not only in semantics but also in related disciplines, especially syntax and the psychology of language and reasoning.

Pronouns and their antecedents
What is the relationship between a pronoun and its antecedent? Sometimes, as in (1a), the link is a syntactic one; but in cases like (1b) it is not.

(1) a. Wilma$_i$ is writing a poem about herself$_i$.
     b. Barney is writing a book$_i$. It$_i$ is about Fred.

As usual, the coindexing serves to indicate that the first expression acts as an 'antecedent' to the pronoun, and a much-discussed question in semantics is what that might mean in examples like (1b).

Non-lexical semantics
Ever since Frege, semanticists have been concerned with the problem how the meaning of a compound expression relates to the meanings of its parts, and some (not I) would go so far as claiming that it is *the* problem in semantics.

I have made a somewhat arbitrary selection from what practically everybody would agree are central issues in semantics. The point of this little exercise is the following. Unless I have completely misunderstood Durst's exposition, NSM theory entails that these are all non-issues. They are either futile or else the ongoing discussions about them are. I take it that this consequence is absurd, or very nearly so, but it may not be immediately obvious how it follows from NSM theory. So let me explain how it does.

In the last twenty years there has emerged a general framework for studying the meanings of quantified expressions. The fundamental notion in this framework is that of a generalised quantifier (GQ), which is simply a relation between sets. Within the GQ framework, the core meaning of 'all' (say) is captured by saying that 'all' denotes the subset relation: 'All A are B' is true if and only if $A \subseteq B$. This style of analysis proves to be a fruitful and intuitively attractive way of studying the interpretation of quantified expressions. To illustrate, if quantifiers are construed as relations between sets, we can say that some of them are symmetric while others are not, and symmetry is a property that turns out to be implicated in various linguistic phenomena as well as language processing.

According to NSM theory, quantifiers are primitives, or at least some of them are. According to Durst, 'one', 'two', 'many', 'much', 'some', and 'all' are considered to be primitive in the more recent versions of NSM theory; which is to say that any attempt at articulating the meanings of such expressions is doomed. This consequence is disturbing enough as it is, but to make things worse, NSM is not even in a position of *addressing*, let alone solving, the kind of issues GQ theory has been trained on.

There is an ongoing discussion in the semantic literature about how sequences like the following are interpreted:

(2)   Fred is writing a book$_i$, too. It$_i$ is about halitosis.

There are very different views on the connection between the pronoun in the second sentence and its antecedent in the first. There is one school of thought which maintains that, although the pronoun is not bound in the syntactic sense of the word, there is a sense (which can be made quite precise) in which it is bound by the indefinite. On this view, the pronoun is not unlike a logical variable, though the underlying logic is non-standard. There is another school that sees the anaphoric link in quite different terms. According to it, the pronoun is really a description in disguise, something akin to 'the book Fred is writing'. On this view, the indefinite furnishes descriptive material for the pronoun's hidden content; the pronoun isn't bound in any way.

If Durst is correct, none of this makes any sense. The literature about anaphora is replete with technical terms that are gibberish to the average taxpayer, and therefore it is so much as useless. What is more, there will be

nothing to replace it, because NSM theory lacks the conceptual means for reformulating the main issues in the first place.

Finally, consider the interpretation of compound expressions like 'pink pumpkin'. Many people working in semantics have been concerned with the issue how the meaning of such a compound is affected by the meanings of its parts and the way they are put together. This seems a perfectly sound question to me. But in the NSM framework it cannot even be stated, as far as I can tell. Furthermore, if we want to answer it, we can hardly do without such concepts as 'adjective', 'noun' — not to mention 'meaning'. These are not semantic primitives, I take it, nor are they definable in terms of the kind of primitives recognised by NSM theorists.

If a scientific theory entails that *all* its competitors are fighting windmills, then this is strong evidence that the theory itself is on the wrong track. My impression is that NSM is such a theory.

It should be clear by now that my feelings towards the NSM project are not too warm. Indeed, the project strikes me as decidedly odd. NSM has only a single analytical tool to offer, that of 'reductive paraphrase': the meaning of any expression in any language is to be given by paraphrasing it in any language, not necessarily different from the object language, using only a handful of words, which supposedly express 'universal semantic primitives'. It is glaringly obvious that this procedure is circular, but apart from that it raises more questions than I can list here. For example, what motivates the requirement that semantic primitives must be expressible in all natural languages? There are quite a few candidate primitives that are eminently plausible from a psychological point of view but are not lexicalised in some, or even all, languages. For example, Durst rejects 'verticality' as a semantic primitive on the grounds that it does not belong to 'the basic vocabulary of language users' (the quote quotes Wierzbicka). But there is a lot of psychological evidence that verticality is a fundamental category of cognition. Apparently, that a concept is central to our thinking does not entail that we have a word for it.

Another way of bringing out the flaws in the procedure of reductive paraphrase is that is based on a misguided notion of semantic methodology. NSM spurns the use of technical terms because they would render semantic analyses unintelligible to laymen, and if an informant doesn't understand our analysis, so the argument goes, he cannot confirm or disconfirm it either. This is a *non sequitur* if there ever was one. After all, we don't expect quarks and viruses to understand our theories about *them*, either. Frankly, *qua* semanticist, I just don't care that my mother-in-law fails to grasp my theories, and it does not prevent me from testing my predictions against her intuitions as a native speaker.

Yet another way of getting at the inherent limitations in the NSM pro-

gramme is this. According to Durst and his school, semanticists are in the business of defining meanings, and meanings are, first and foremost, word meanings. That is to say, the project and methods of the semanticist are those of the lexicographer. But no matter how central the lexicon may be, and how tenuous the distinction between lexicon and grammar or between meaning and use, the project of reducing all of semantics to lexicography is just a pie in the sky.

The authors of the Longman dictionary of English have managed to cast all their definitions in about 2000 words, and that was an extraordinary feat. Wierzbicka and her following believe that less than 20 words will suffice for a full-blown semantics of any language. That sums it up rather nicely, I should think.