Words and Worlds: On the linguistic analysis of modality

Richard Matthews. Frankfurt: Peter Lang. 1991. Vol. 191 of the European University Series XIV, Anglo-Saxon Language and Literature. 310 pp., including 13 tables and two figures.

This book is particularly valuable for the way in which it conducts investigation in descriptive linguistics—namely with an eye toward how the description will fit in a more general theory of language. This welcome rapprochement of descriptive and theoretical work is particularly interesting to me as a practitioner in natural language processing (NLP), an area of applied linguistics.

NLP involves an application of linguistic knowledge to the problems of processing speech and text on the computer. After an early phase in which the field tried to do without linguistic expertise altogether, NLP today embraces linguistics for morphology, syntax and semantics. NLP is today in this sense applied linguistics. (This should not be taken to mean that it is not applied computer science, which it also is.)

But the linguistics being applied is THEORETICAL, not DESCRIPTIVE. Given the difficulty of programming even the relatively exact linguistic work ones finds say in *Generalized Phrase Structure Grammar* (Gazdar et al. 1985) or *Lexical Functional Grammar* (Bresnan 1982), it was necessary for the field to use linguistic theory as a start. But now that the question of how to implement these theories is fairly well understood (though improvements are still the subject of very active research), the field is running up against a roadblock of knowledge acquisition: if it takes a linguist a year to read up on and code the information in 200 words—when can we have linguistically sophisticated systems with vocabularies of 20,000 words (and four times as many forms)? One might think that having 100 people work on the problem for a year could help, but this underestimates the difficulty of coordinating such sensitive work. There is no chance that the resulting system would be consistent, let alone well-organized or optimized.

NLP researchers are now very actively trying to gather data AUTOMATICALLY, either through the statistically guided analysis of large (10^8 words) corpora, or through the careful analysis of dictionaries. But these efforts remain in their infancy: the highly articulated information needed for exact lexical description in the theories in current NLP use is too subtle for current corpus tools.

The field of DESCRIPTIVE LINGUISTICS, especially that branch which tries to inform and be informed by theoretical linguistics, could play a most important role here. If NLP could even partially acquire the lexical knowledge it needs from works in descriptive linguistics, then it could partially solve the problem of the data bottleneck.

Richard Matthews has written a useful and timely book which—while primarily descriptive—insists on the necessary dialogue between theory and description. The subject matter is most complex, the semantics of modality, and Matthews is assiduous in collecting and organising his descriptive material. He examines not only modal verbs but also a range of semantically similar constructions, such as *X is thought to ...*, *is supposed to ...*, etc.

The choice of data and methodology is of course crucial for a work on semantics as subtle as that of modality, and Matthews has settled on an intelligent compromise. On the one hand he has attempted to find collocational tests which are sensitive to meaning, so that his test data ultimately have the same status as acceptability data in syntax or morphology. This ensures replicability and perhaps corpus testing in the future. On the other hand, however, he frequently includes raw reports of semantic intuition in his data summaries (e.g., whether two forms are semantically equivalent, whether a form can be understood as a request or command, or whether a form is understood as reporting a state of affairs "confidently"). This sort of information could be useful to the theorist, but also, I suppose, to the applications specialist in language pedagogy.

The work will remain useful for years to come if only for the enormous amount of data Matthews assembles here. Pages and pages of acceptability judgements are carefully organized, and the tables and figures each summarize hundreds of data points involving acceptability and semantic properties. I believe that the data is reported reliably, but it is not surprising that some of Matthews's data annotations do not quite jibe with the prejudices of an American ear (e.g., *He probably ought to have arrived by now*, p.67, deserves no '?" in my mind), but this is to be expected, especially in view of the thousands of data points recorded here.

The content of the book is innovative in proposing an extension of the usual analytical and classificatory apparatus used in modal semantics. The primary thesis of the work is that modal elements contribute semantically to several different dimensions of meaning. He begins the work by isolating three: the illocutionary, the modal proper, and the propositional nucleus. To this he adds a temporal dimension in the course of the investigation.

illocutionary aspects of meaning concern communicative function—asserting, denying, requesting, etc. Thus the modal verb *must* can be used to express a command, an illocution.

modal aspects of meaning indicate the speakers attitude toward the propositional nucleus—irrealis, factive, potential. The modal verb *must* may express the (epistemic) necessity, a modality.

the propositional nucleus is the content of what is said less illocution and modality—the event or state of affairs related. Thus the modal verb *can* is often used simply to report on an ability, as in *He can swim.*

temporal aspects of meaning distinguish several modal expressions minimally, and this dimension is therefore added for completeness.

These dimensions are not on equal footing. Illocutionary meaning relates all of the other aspects to communicative function, while modal meaning proper relates propositional nucleus—including temporal meaning—to various "worlds" of the title. Matthews therefore suggests viewing these as functions applying to one another's output:

$$U \ : \ \mathbf{Ill(Mod(Nuc))}$$

A limited amount of recursion is admitted by allowing **Nuc** to contain elements from **Mod**. When proposing a classification of modals, Matthews generally provides three (or four) independent specifications of the content associated with the modal, which seems sufficient.

The specifications associated with **Mod** may involve not only times but also 'possible worlds'—e.g., worlds in which a given propositional obtains (or fails to obtain), or an irreal world. One would assume that Matthews would proceed from these concepts to the natural framework in which to study the modes of a proposition holding, that of modal logic, but Matthews resists this step for reasons I find ultimately uncompelling. Matthews observes quite correctly (p.27) that, in many of the sorts of "necessity" concepts employed in natural language, a proposition p is not entailed by a proposition $\Box p$ 'it is necessary that p. This entailment relation obtains in those systems employing the axiom **T**:

$$\mathbf{T} \quad \Box p \supset p$$

But standard texts on modal logic introduce the subject via "normal" logics **K** in which the axiom need not hold. Cf. G.E.Hughes and M.J.Cresswell 1968 or Chellas 1980. Deontic logic, in which the necessity operator is interpreted as 'it ought to be the case that' is a particularly clear case of a logic which eschews **T**. Matthews goes on to argue that English "does not normally employ expression in assertions, which arguably, involve logical necessity, such as: *A bachelor is unmarried.*" But this seems to introduce an implicit assumption that certain uses of the copula could not involve an semantic modal. We might reject the implicit assumption, however, or, alternatively, we might allow that other expressions of necessity be admitted into a semantic representation scheme.

Finally, Matthews alludes to a "fundamental problem" (p.28), namely that "the closed internally consistent systems set up for modal logic [...] cannot be taken as the terms of a semantic specification of open, non-consistent and, in fact, polysystemic modality of natural language." It would exceed the bounds of a review to try to answer these indeed fundamental objections here, but I do not believe that such fears need to remain with us forever. What I believe is correct here is that there are no systems which anyone might employ "off the shelf" for the purpose of providing the analyses needed for all the range of modality analysed here. But the attempt to identify a modal logic system modeling (part of) Matthews's data might have systematized further, sharpened further questions on data, or even suggested refinements of the analytic and classificatory system.

The technical aspects of the book are satisfactory. The book is carefully edited and quite nicely typeset. A very detailed "Guide to Contents" compensates somewhat for the regrettable lack of an index.

I recommend *Words and Worlds* to researchers in descriptive and theoretical linguistics, especially those interested in the cooperation of the subdisciplines and most especially to students of modal semantics.

References

Bresnan, J. (ed.). 1982. *The Mental Representation of Grammatical Relations.* Cambridge, Mass.: MIT Press.

Chellas, B. F. 1980. *Modal Logic: An Introduction.* Cambridge: Cambridge University Press.

Gazdar, G., E. Klein, G. Pullum, and I. Sag. 1985. *Generalized Phrase Structure Grammar.* Harvard University Press.

G.E.Hughes, and M.J.Cresswell. 1968. *An Introduction to Modal Logic.* London: University Paperbacks.

John Nerbonne
nerbonne@let.rug.nl

Alfa Informatica
Postbus 716
Oude Kijk in 't Jatstraat 41
Rijkuniversiteit Groningen
NL 9700 AS Groningen
The Netherlands