Mass terms and model-theoretic semantics
In this series

1. David Crystal: *Prosodic systems and intonation in English*
2. John M. Anderson: *The grammar of case*
3. M. L. Samuels: *Linguistic evolution*
4. P. H. Matthews: *Inflectional morphology*
5. Gillian Brown: *Phonological rules and dialect variation*
6. Brian Newton: *The generative interpretation of dialect*
7. R. M. W. Dixon: *The Dyirbal language of North Queensland*
8. Melissa Bowler: *Early syntactic development*
9. W. Sidney Allen: *Accent and rhythm*
10. Peter Trudgill: *The social differentiation of English in Norwich*
11. Roger Lass and John M. Anderson: *Old English phonology*
12. Ruth M. Kempson: *Presupposition and the delimitation of semantics*
13. James R. Hupford: *The linguistic theory of numerals*
14. Roger Lass: *English phonology and phonological theory*
15. G. M. Awbery: *The syntax of Welsh*
16. R. M. Dixon: *A grammar of Yidi*
17. James Foley: *Foundations of theoretical phonology*
18. A. Radford: *Italian syntax: transformational and relational grammar*
19. Dieter Wunderlich: *Foundations of linguistics*
20. David W. Lightfoot: *Principles of diachronic syntax*
21. Annette Karmiloff-Smith: *A functional approach to child language*
22. Per Linell: *Psychological reality in phonology*
23. Christine Tanz: *Studies in the acquisition of deictic terms*
24. Roger Lass: *On explaining language change*
25. Torben Thrané: *Referential-semantic analysis*
26. Tamsin Donaldson: *Nyguambaa*
27. Kristján Árnason: *Quantity in historical phonology*
28. John Laver: *The phonetic description of voice quality*
29. Peter Austin: *A grammar of Diyarri, South Australia*
30. Alice C. Harris: *Georgian syntax*
31. Suzanne Romaine: *Socio-historical linguistics*
32. Martin Atkinson: *Explanations in the study of child language development*
33. Suzanne Fleischman: *The future in thought and language*
34. Jenny Cheshire: *Variation in an English dialect*
35. William A. Foley and Robert D. van Valin Jr: *Functional syntax and universal grammar*
36. Michael A. Covington: *Syntactic theory in the High Middle Ages*
37. Kenneth J. Safir: *Syntactic chains*
38. J. Miller: *Semantics and syntax*
39. Harry C. Bunt: *Mass terms and model-theoretic semantics*
40. Heinz J. Giegerich: *Metrical phonology and phonological structure*

Supplementary Volumes

Brian D. Joseph: *The synchrony and diachrony of the Balkan infinitive*

* Issued in hard covers and as a paperback
MASS TERMS AND MODEL-THEORETIC SEMANTICS

HARRY C. BUNT
Department of Language and Literature,
Tilburg University

CAMBRIDGE UNIVERSITY PRESS
CAMBRIDGE
LONDON NEW YORK NEW ROCHELLE
MELBOURNE SYDNEY
Contents

Preface ix

PART 1
MODEL-THEORETIC SEMANTICS OF MASS TERMS

1 Introduction 3
1.1 Semantic issues relating to mass terms 3
1.2 Plan of the book 6

2 Mass terms 9
2.1 Mass terms syntactically defined 9
2.2 Mass terms semantically defined 15

3 Approaches to mass term semantics 21
3.1 Differences in approaches 21
3.2 Quine’s semi-mereological approach 22
3.3 Fully mereological approaches 26
3.4 Mass nouns as denoting substances 29
3.5 Set-based approaches 33

4 Towards a semantic theory of mass nouns 43
4.1 Requirements for a semantic theory of mass nouns 43
4.2 The minimal parts hypothesis 44
4.3 The use of mereology 47

5 Ensemble theory 53
5.1 Part–whole and member–whole 53
5.2 The part–whole relation 56
5.3 Atoms, unicles, and members 59
5.4 Continuous and discrete ensembles 63
5.5 A model for ensemble theory 66
5.6 Ensembles, sets, and individuals 71
vi  Contents

6  Semantic representations based on ensemble theory 73
6.1  Introduction 73
6.2  Amount terms 75
6.3  Language definition and families of languages 81
6.4  The EL type system 84
6.5  The syntax of EL 88
6.6  The semantics of EL 92
6.7  Semantic representations in EL 94
6.8  Notation: summary and some additions 99

7  Two-level model-theoretic semantics 102
7.1  Formal semantics 102
7.2  Model-theoretic semantics 105
7.3  Model-theoretic analysis of natural language 111
7.4  Referential ambiguity and model-theoretic semantics 114
7.5  Two-level model-theoretic semantics 118
7.6  A formal and a referential view of the world 124
7.7  Mass terms in two-level model-theoretic semantics 129

8  Quantification and mass nouns 133
8.1  Introduction 133
8.2  Quantification 137
8.3  Quantification aspects of multi-NP sentences 142
8.4  Quantification representations for count nouns 146
8.5  Quantification in two-level model-theoretic semantics 152
8.6  Mass noun quantification 161
8.7  The representation of mass noun quantification 165
8.8  A grammar for a fragment of English 180

9  Modification and mass nouns 197
9.1  Modification and quantification 197
9.2  Theoretical problems with adjective–mass noun combinations 199
9.3  Empirical facts about adjective–mass noun combinations 204
9.4  The representation of mass noun modification 212
9.5  Augmented phrase-construction grammar for an extended fragment of English 217
PART 2  
ENSEMBLE THEORY

10 Axiomatic ensemble theory 233
10.1 The logical system 233
10.2 Ensemble theory built up from a system of axioms 235
10.2.1 Axiom of transitivity 235
10.2.2 Axiom of equality 236
10.2.3 Axiom of unicles 237
10.2.4 Axiom of extensionality 238
10.2.5 Axiom of pairing 242
10.2.6 Axiom of powers 244
10.2.7 Axiom of merging (or sums) 244
10.2.8 Axiom (schema) of replacement 247
10.2.9 Axiom of infinity 251
10.2.10 Axiom of regularity 252

10.3 Operations on ensembles 252
10.3.1 Merge 252
10.3.2 Overlap 252
10.3.3 Completion 254
10.3.4 The partition theorem 255
10.3.5 The Boolean algebra of ensemble operations 256

11 Continuous, discrete, and mixed ensembles 260
11.1 Continuous ensembles 260
11.2 Discrete ensembles 263
11.3 Continuity versus discreteness 266

12 A model for ensemble theory 269
12.1 The model structure 270
12.2 Satisfaction of the ETX axioms 271

13 Ensemble theory, set theory, and mereology 275
13.1 Ensemble theory and set theory 275
13.1.1 Sets in ensemble theory 278
13.1.2 Ensembles in set theory 281
13.2 Ensemble theory and mereology 288
13.2.1 Mereology, or the theory LLG 288
Contents

13.2.2 Ensembles and individuals 290
13.2.3 LLG individuals in ensemble theory 295
13.2.4 Ensembles in the theory LLG 297

Notes 303
Bibliography 313
Index 319
Preface

This book has grown out of my dissertation, submitted to the University of Amsterdam in 1981. It contains most of the ideas of the dissertation in a revised form, but it also contains a good deal of new material.

New material, that supplements the dissertation at an important point, concerns the description of a model of the axioms of ensemble theory, the generalization of set theory designed for the semantic analysis of mass terms. Section 5.6 contains an informal presentation of a model, and chapter 12 is devoted to the description of a class of models in formal terms. New material relating to ensemble theory can also be found in the final chapter, where a detailed discussion is given of the relation between ensemble theory, classical set theory, and mereology (also known as ‘calculus of individuals’). This should prevent such misunderstandings as that ensemble theory is ‘another name for mereology’ (Roeper, 1983).

Other new material concerns the definition of the general framework in which the semantic analysis is carried out. In this framework two stages of analysis are distinguished: one where the logical semantic aspects of a natural language expression are analysed and one where the referential semantic aspects are analysed in relation to a model of a certain domain of discourse. This has the advantage that the analysis of general, domain-of-discourse independent semantic aspects and that of purely lexical semantic aspects are well-separated components of a single overall framework. Each stage of the analysis results in a representation in a formal language (more than one representation in the case of ambiguities). This gives rise to an interesting treatment of lexical ambiguity and vagueness: content words in the natural language are represented at the nonreferential level by constants of the representation language that preserve the ambiguity and vagueness of the corresponding natural language words; the precise meanings of these words are established by interpretation rules that have the effect of ‘translating’ these constants into expressions of a second formal language, whose constants have a one-to-one correspondence to entities in the model of the
Preface

discourse domain. A precise definition of this framework is given in chapter 7 in the terminology of model-theoretic semantics. The relation between the two representation languages, mediating between the natural language and the discourse-domain model, is defined by model-theoretic interpretation rules, where the expressions of the representation language at the referential level serve as interpretations of those of the other representation language. The framework is therefore called ‘two-level model-theoretic semantics’.

The most important new material in this book concerns the treatment of ambiguity and vagueness in quantification. The chapter on quantification in my dissertation was primarily meant to illustrate the power of ensemble theory in the analysis of mass noun expressions. As such it may have served its purpose. But it seemed intuitively obvious that the two-level framework had not been fully exploited. The new idea which underlies the treatment of quantification in this book is that the separation of domain-of-discourse dependent meaning aspects from those that do not depend on the domain of discourse is applied to other areas of semantic investigation than that of lexical semantics. For example, the sentence ‘The boys carried two pianos upstairs’ may have several, quantificationally different readings: one where each of the boys carried each of the pianos, one where the boys collectively carried each of the pianos, one where each of the boys carried a collection of two pianos, etc. All these interpretations are formally possible, but usually not all of them are possible in a given domain of discourse. That depends on whether one considers 18-year-old boys or eight-year-old boys, real pianos or toy pianos, and other such domain-of-discourse dependent considerations. It is therefore attractive to leave such decisions as to whether the boys acted collectively, individually, or in groups to the domain-of-discourse dependent part of the analysis. I have devised a way of doing this, which entails that at the purely formal level of analysis a quantified sentence has a representation which is ambiguous or vague in those quantification aspects on which the natural language sentence is not specific. A sentence like ‘The boys carried two pianos upstairs’ has only one representation at this level, and the number of interpretations it has at the referential level is determined by what the model of the discourse domain says about boys, pianos, and carrying something upstairs. This makes it possible to formulate an articulate analysis of quantifications while at the same time preventing an uncontrolled explosion of interpretations.

The analysis of constructions with adnominal modification, such as ‘The pianos carried upstairs by the boys’, entails many of the same considerations
as that of quantified expressions. I have therefore also devised a new
treatment of the ambiguity and vagueness in adnominal modification along
the same lines as the treatment of quantification. An interesting minor
point which is new in the chapter on modification is the treatment of
'collective modification'. This occurs for instance in the sentence 'John has
heavy books in his bag' on the interpretation that the collection of books
in the bag as a whole is heavy. This case does not easily fit into a systematic
syntactic/semantic analysis, since the adjective 'heavy' is syntactically a
modifier of 'books', while semantically it applies to 'the books John has in
his bag'.

A new form has also been given to the syntactic/semantic grammar rules,
used in the chapters on quantification and modification to describe frag-
ments of English. There were two independent reasons to reconsider the
form of the grammar. An anonymous referee for Cambridge University
Press had complained about the original form, which had been borrowed
from Scha (1981), and asked for a formulation that would be more in line
with currently common grammatical formalisms. Moreover, work on a
computer implementation of the grammar made it clear that certain changes
in the rules would permit more transparent parsing algorithms. The changes
that have been made are: (1) the syntactic component of a rule has been
split up into three parts: (a) a constituent-structure rule, (b) a condition
on syntactic features, and (c) a feature-propagation rule; (2) a form of
constituent-structure rule has been designed which is not purely concatenat-
ing and which allows a restricted form of context-sensitivity. This grammar,
for which I have coined the term 'Augmented phrase-formation gram-
mar', belongs to the class of phrase-structure-based grammar formalisms
that includes currently developing formalisms such as Generalized phrase-
structure grammar (Gazdar, 1982; Gazdar and Pullum, 1982) and Aug-

As a study in semantics, this book should in the first place be of interest
to linguists and language philosophers. But since the study raises issues
concerning the logical and set-theoretical foundations of model-theoretic
semantics, it also aims at an audience of logicians and mathematicians. This
motivates the division of this book into two parts. Moreover, the semantic
analysis of mass terms is partly based on arguments relating to why people
express something in one way rather than another, with special attention
to the fact that the logical structure of a speaker's apparent beliefs as
expressed by his words should be distinguished from the logical structure
xii Preface

of his actual beliefs. These points, and their far-reaching consequences for the general framework of semantic analysis, may be of interest to psycholinguists. Finally, this study was directly motivated by the problems that arose in an attempt to design a computer program for syntactic and semantic analysis of natural language. The solutions which are presented here, and which are made explicit in the form of grammar rules and interpretation rules for fragments of English, may be of interest to computer scientists and researchers in Artificial Intelligence, who may also be interested in the possible use of ensemble theory in knowledge representation in intelligent systems (see Bunt, 1985). It is because of this wide variety of potential readers that most of the chapters in this book have an introduction at a rather elementary level. Chapter 6, for instance, which discusses the use of ensemble concepts in semantic representations, contains introductory sections on the nature of semantic representations in general and on the definition of representational formalisms. Similarly, chapter 7 contains a general introduction to model-theoretic semantics and ambiguity. My hope is that these sections will enable readers from a variety of academic disciplines to understand the technically more complex material.

I should like to thank a number of people who have contributed directly or indirectly to the realization of this book. In the first place I want to thank my former colleagues at Philips Research Laboratories in Eindhoven for years of fruitful cooperation in the PHILQA project, where many of the basic ideas in this book were born. My thanks go to Wim Bronnenberg, Jan Landsbergen, Piet Medema, Remko Scha, Wijnand Schoenmakers, and Eric van Utteren. I also want to thank my colleagues and students at the Institute for Perception Research (IPO) in Eindhoven for participating in discussions which helped to shape some of the ideas in this book, for engaging in computer implementation of semantic analysis of fragments of Dutch and English using the two-level framework and the grammar formalism developed here (thereby debugging the corresponding sections in the book), and in general for providing a very pleasant and stimulating working environment. My thanks go in particular to Herman Bouma, Don Bouwhuis, the late Ab van Katwijk, Jurgen van der Linden, Herman Muller, Floris van Nes, Robbert-Jan Beun, Frens Dols and Gemme thoe Schwartzenberg. I feel very fortunate to be in the position to continue my research in the new linguistics department at Tilburg University, in close cooperation with IPO. I also want to thank those with whom I have had fruitful discussions about my dissertation or new ideas incorporated in this book, including
my thesis supervisors Hugo Brandt Corstius and Simon Dik; Jens Allwood, Johan van Benthem, Peter van Emde Boas, Jeroen Groenendijk, Sjaak de Mey, Julius Moravesik, Jane Robinson, and Remko Scha. For comments on some version of the manuscript of this book I thank the anonymous referees of Reidel and Cambridge University Press, Robbert-Jan Beun, Dolf van Wezel and Geraldine Stoneham.

Finally, and most importantly, I want to thank my wife Annette for her careful proof reading of the manuscript and for helping me to bring it to its final form, and in general for helping me to keep going. Without her continuous support this book would not have been written.