Count and Mass Across Languages
Oxford Studies in Theoretical Linguistics

General Editors
David Adger and Hagit Borer, Queen Mary, University of London

Advisory Editors
Stephen Anderson, Yale University; Daniel Büring, University of California, Los Angeles; Nomi Erteschik-Shir, Ben-Gurion University; Donka Farkas, University of California, Santa Cruz; Angelika Kratzer, University of Massachusetts, Amherst; Andrew Nevins, University College London; Christopher Potts, Stanford University, Amherst; Barry Schein, University of Southern California; Peter Svenonius, University of Tromsø; Moira Yip, University College London

Recent titles
24 Quantification, Definiteness, and Nominalization
   edited by Anastasia Giannakidou and Monika Rathert
25 The Syntax of Sentential Stress
   by Arsalan Kahnemuyipour
26 Tense, Aspect, and Indexicality
   by James Higginbotham
27 Lexical Semantics, Syntax, and Event Structure
   edited by Malka Rappaport Hovav, Edit Doron, and Ivy Sichel
28 About the Speaker
   Towards a Syntax of Indexicality
   by Alessandra Giorgi
29 The Sound Patterns of Syntax
   edited by Nomi Erteschik-Shir and Lisa Rochman
30 The Complementizer Phase
   edited by E. Phoevos Panagiotidis
31 Interfaces in Linguistics
   New Research Perspectives
   edited by Raffaella Folli and ChristianeUlbrich
32 Negative Indefinites
   by Doris Penka
33 Events, Phrases, and Questions
   by Robert Truswell
34 Dissolving Binding Theory
   by Johan Rooryck and Guido Vanden Wyngaerdt
35 The Logic of Pronominal Resumption
   by Ash Asudeh
36 Modals and Conditionals
   by Angelika Kratzer
37 The Theta System
   Argument Structure at the Interface
   edited by Martin Everaert, Marijana Marelj, and Tal Siloni
38 Sluicing
   Cross-Linguistic Perspectives
   edited by Jason Merchant and Andrew Simpson
39 Telicity, Change, and State
   A Cross-Categorial View of Event Structure
   edited by Violeta Demonte and Louise McNally
40 Ways of Structure Building
   edited by Myriam Uribe-Etxebarria and Vidal Valmala
41 The Morphology and Phonology of Exponence
   edited by Jochen Trommer
42 Count and Mass Across Languages
   edited by Diane Massam
43 Genericity
   edited by Alda Mari, Claire Beyssade, and Fabio Del Prete
44 Strategies of Quantification
   edited by Kook-Hee Gil, Steve Harlow, and George Tsoulas

For a complete list of titles published and in preparation for the series, see pp. 311–12.
Count and Mass Across Languages

Edited by
DIANE MASSAM
Contents

General Preface ix
The Contributors x
Abbreviations xv

1. The count mass distinction: Issues and perspectives 1
   Jila Ghomeshi and Diane Massam

   1.1 Setting the stage 1
   1.2 Are count and mass conceptually universal and are they mapped
to the real world uniformly? 3
   1.3 Are count and mass universally expressed in language and are
they always expressed in the same way? 4
   1.4 If count and mass are not expressed in the same way universally,
how are they expressed? 5
   1.5 Conclusion 7

2. Lexical nouns are both +mass and +count, but they
   are neither +mass nor +count 9
   Francis Jeffry Pelletier

   2.1 Introduction: Informal accounts of +mass and +count 9
   2.2 +mass and +count as syntax 11
   2.3 +mass and +count as semantics 11
   2.4 Problems with the syntactic approach 13
   2.5 Problems with the semantic approach 15
   2.6 Evaluation, and a flaw in common 17
   2.7 A different approach 17
   2.8 Related proposals 21
   2.9 Cross-linguistic comments 23
   2.10 Further advantages 24
   2.11 A final philosophical remark 26

3. Aspects of individuation 27
   Elizabeth Cowper and Daniel Currie Hall

   3.1 Introduction 27
   3.2 Taxonomy of English nouns 30
   3.3 Chinese 34
   3.4 Plurality and classifiers 36
Contents

3.5 Plural marking in Chinese? 37
3.6 Apparent plural marking in Korean 42
3.7 Individuation in Persian 45
3.8 Conclusions 52

4. Collectives in the intersection of mass and count nouns:
   A cross-linguistic account 54
   Heike Wiese
   4.1 Introduction 54
   4.2 Conceptual and morphosyntactic distinctions in the mass/count domain 57
   4.3 Variation in syntactic-conceptual mass/count correspondences 66
   4.4 Semantics as a mediator of syntactic and conceptual classifications 71
   4.5 Conclusions 73

5. Individuation and inverse number marking in Dagaare 75
   Scott Grimm
   5.1 Introduction 75
   5.2 The semantic basis of inverse number marking in Dagaare 77
   5.3 Language internal correlates 87
   5.4 Cross-linguistic correlates 90
   5.5 A formal account of -ri 94
   5.6 Conclusion 97

6. General number and the structure of DP 99
   Ileana Paul
   6.1 Introduction 99
   6.2 Malagasy 100
   6.3 Cross-linguistic considerations 107
   6.4 The emerging typology 110

7. Plural marking beyond count nouns 112
   Saeed Ghaniabadi
   7.1 Introduction 112
   7.2 Background on Persian noun phrases 113
   7.3 Data 114
   7.4 Definiteness/Number syncretism 121
   7.5 Categorial identity of Persian plural marker 122
   7.6 Analysis 124
   7.7 Conclusion 128
8. Aspeclual effects of a pluractional suffix: Evidence from Lithuanian  
*Solveiga Armoskaite*  
8.1 The problem: -inė- is not a dedicated aspectual morpheme  
8.2 Suffix -inė- is Number  
8.3 The aspeclual effects of -inė- are epiphenomenal  
8.4 Conclusions & further questions  

9. Decomposing the mass/count distinction: Evidence from languages that lack it  
*Martina Wilschko*  
9.1 Introduction  
9.2 The mass/count distinction is not universally associated with categorical properties  
9.3 The source of the categorical properties of the mass/count distinction  
9.4 Variation in the content of the categorical properties:  
[± bounded] versus [± animate]  
9.5 Conclusion  

10. On the mass/count distinction in Ojibwe  
*Eric Mathieu*  
10.1 Introduction  
10.2 Number as an inflectional category in Ojibwe  
10.3 Ojibwe pluralized mass nouns  
10.4 Basis for a solution  
10.5 Conclusion  

11. Counting and classifiers  
*Lisa Lai-Shen Cheng*  
11.1 Introduction: Three puzzles  
11.2 Count/mass at the classifier level  
11.3 Chinese classifiers  
11.4 Do all classifiers individuate or divide?  
11.5 Plural classifiers  
11.6 Conclusion  

12. Countability and numeral classifiers in Mandarin Chinese  
*Niina Ning Zhang*  
12.1 Introduction  
12.2 Decomposing countability
12.3 The two features in nouns 227
12.4 The two features in unit words 229
12.5 Comparing with the dichotomous-contrast analysis 231
12.6 Reflections on theories of the relation between CLs and countability 233
12.7 Summary 237

13. Semantic triggers, linguistic variation and the mass-count distinction 238
   Alan C. Bale and David Barner
   13.1 Introduction 238
   13.2 English, Mandarin, and the mass-count distinction 239
   13.3 English and Mandarin heuristics and the age of distinction 248
   13.4 Possible triggers for a parametric distinction 251
   13.5 Conclusion 259

14. Classifying and massifying incrementally in Chinese language comprehension 261
   Natalie M. Klein, Greg N. Carlson, Renjie Li, T. Florian Jaeger, and Michael K. Tanenhaus
   14.1 Background 261
   14.2 Previous research 263
   14.3 Experiment One: English measure phrases 265
   14.4 Experiment Two: Chinese massifiers 271
   14.5 Experiment Three: Chinese count classifiers 275
   14.6 Comparing across language and ontology 278
   14.6 Conclusions 282

References 283
Index 305
General Preface

The theoretical focus of this series is on the interfaces between subcomponents of the human grammatical system and the closely related area of the interfaces between the different subdisciplines of linguistics. The notion of ‘interface’ has become central in grammatical theory (for instance, in Chomsky’s recent Minimalist Program) and in linguistic practice: work on the interfaces between syntax and semantics, syntax and morphology, phonology and phonetics, etc. has led to a deeper understanding of particular linguistic phenomena and of the architecture of the linguistic component of the mind/brain.

The series covers interfaces between core components of grammar, including syntax/morphology, syntax/semantics, syntax/phonology, syntax/pragmatics, morphology/phonology, phonology/phonetics, phonetics/speech processing, semantics/pragmatics, and intonation/discourse structure, as well as issues in the way that the systems of grammar involving these interface areas are acquired and deployed in use (including language acquisition, language dysfunction, and language processing). It demonstrates, we hope, that proper understandings of particular linguistic phenomena, languages, language groups, or inter-language variations all require reference to interfaces.

The series is open to work by linguists of all theoretical persuasions and schools of thought. A main requirement is that authors should write so as to be understood by colleagues in related subfields of linguistics and by scholars in cognate disciplines.

The count–mass distinction is, as the introduction to this volume makes clear, a perfect ‘workshop’ for investigating major issues to do with the interaction between language, cognition and the world, and has been central in both linguistic and philosophical discussions over the decades. The present volume brings together scholars and researchers who use cross-linguistic and experimental methods to investigate the extent to which the count–mass distinction is a fundamental linguistic category. The emerging consensus is that this category is indeed fundamental, but that its expression in languages of the world is tightly constrained by grammatical factors and by factors that lie at the interface of syntax and compositional semantics.

David Adger
Hagit Borer
The Contributors

Solveiga Armoskaite is a Visiting Scholar at University of California, Los Angeles. She recently defended her dissertation *The destiny of roots in Blackfoot and Lithuanian* (2011, University of British Columbia, Vancouver). Her current research interests are argument structure, aspect, Blackfoot, categorization, determiners, Heritage Language, Lithuanian, morphosyntax, re-categorization, and roots.

Alan C. Bale (Ph.D. 2006, McGill University) is a semanticist working in Montreal as a Research Associate (Concordia University) and Faculty Lecturer (McGill University). His work focuses on the grammatical representation of scales and quantification and has been published in a variety of journals that span many different subfields of linguistics, including semantics, morphology and language acquisition. He is currently involved in a research project studying number marking in English, French, Western Armenian and Mi’kmaq.

David Barner received his Ph.D. in 2006 from Harvard University. He is an Assistant Professor of Psychology at the University of California, San Diego. His work spans topics including the mass-count distinction, linguistic relativity, the acquisition of nouns and verbs, counting, and pragmatic development. Recently, his work has focused on the relationship between language acquisition and children’s concepts of time, color, and space.

Greg N. Carlson (M.A. University of Iowa, 1974; Ph.D. University of Massachusetts at Amherst, 1977) is a Professor of Linguistics, Philosophy, and Cognitive Sciences at the University of Rochester. His research interests include natural language semantics, focusing on the semantics of generics and bare plurals and related issues. He also conducts research in experimental psycholinguistics, with other strong interests in philosophy and computer science. Much of this research has been supported by NSF and NIH. Publications include *Reference to Kinds in English* (Garland, 1980), ‘Marking constituents’ (1983), *The Generic Book* (Chicago, edited with F. J. Pelletier, 1995), ‘Generic Passages’ (NLS, 1997), ‘The Average American Has 2.3 Children’ (*Journal of Semantics*, 2002), and ‘Weak Definates,’ (2006). He served as editor in chief of *Linguistics and Philosophy* (1992–1997) and of *Language* (2009–).

Lisa Lai-Shen Cheng is Chair Professor of Linguistics at Leiden University. Her research interests include comparative syntax, the syntax-semantics interface and syntax-phonology interface. Her work on comparative syntax includes ‘Bare and not-so-bare nouns and the structure of NP’ (1999, *Linguistic Inquiry*, with Rint Sybesma), and ‘Bantu relatives’ (2006, *NELS 36*). Her recent work on the syntax-

**Elizabeth Cowper** (Ph.D. 1976, Brown University) is Professor of Linguistics at the University of Toronto, where she has taught since 1976. Her research has focused on auxiliary verbs and the functional structure of the clause, and on number, definiteness and the functional structure of the nominal. She has published in the *Canadian Journal of Linguistics, Studia Linguistica, the Linguistic Review, Language*, and *Linguistic Inquiry*, and is the author of a textbook on the theory of Government and Binding, published by the University of Chicago Press.

**Jila Ghomeshi** is an Associate Professor of Linguistics at the University of Manitoba. She has had a long-standing interest in the syntax of noun phrases. She has published journal articles on number, the Ezafe construction in Persian, and reduplication of English nouns. She has worked on proper names with Diane Massam and has co-edited a book on Determiners with Ileana Paul and Martina Wiltshcko. Her language interests are primarily in Persian and English. Apart from her work on syntax, she has also published a short book on prescriptivism for a general audience entitled *Grammar Matters*.

**Scott Grimm** is a Ph.D. student in Linguistics at Stanford University. His primary research interests revolve around semantics and its effects on grammatical expression. He has previously worked on a variety of topics including morphological case, raising/control and statistical modeling of language use. His dissertation explores the count-mass distinction from a cross-linguistic perspective.

**Saeed Ghaniabadi** (Ph.D. 2010, University of Manitoba) is an Assistant Professor of Linguistics at Hakim Sabzevari University, Iran. His research focuses on Persian morphosyntax. His recent projects include mobile plural marking, Q- versus N-partitive constructions, and the empty noun construction in Persian. He has published a book chapter and presented several papers at international conferences.

**Daniel Currie Hall** (Ph.D. 2007, University of Toronto) is an Assistant Professor of Linguistics at Saint Mary’s University in Halifax, Nova Scotia. He has previously taught at the University of Toronto and at Queen’s University, and worked as a researcher at the Meertens Institute of the Royal Netherlands Academy of Arts and Sciences. His research focuses on the organization and interpretation of distinctive features and the contrasts that they mark, in both phonology and morphosyntax.

**T. Florian Jaeger** is the Wilmot Assistant Professor of the Sciences at the Department of Brain and Cognitive Sciences, University of Rochester. His work focuses on how the computational systems underlying language acquisition, processing, and
production facilitate robust and efficient communication. Recent and ongoing research in his lab includes work on information theoretic and Bayesian accounts of efficient language production (Frank and Jaeger, 2008; Jaeger, 2010, 2011), uncertainty reduction in language acquisition (Fedzechkina, Jaeger, and Newport, 2011, in progress), and rapid adaptation to changed situational statistics in phonetic and syntactic processing (Jaeger and Snider, 2008, submitted; Fine, Jaeger, Farmer, and Qian, in prep; Kleinschmidt and Jaeger, 2011). Jaeger has also contributed to the development of statistical methods, in particular for behavioral and typological work (Jaeger, 2008, in prep; Jaeger, Graff, Croft, and Pontillo, 2011; Jaeger and Pontillo, submitted). His research is currently sponsored by an Alfred P. Sloan Research Fellowship and by NSF grants.

Natalie Klein is a recent graduate from the Department of Brain and Cognitive Sciences at the University of Rochester in Western New York. She is a graduate of Pomona College, where she studied Linguistics and Cognitive Science and minored in Chinese. She is interested in the relationship between world-based knowledge, language, and mental representation, and her doctoral research explores this intersection by using experimental approaches to understanding weak definite noun phrases.

Renjie Li attended the University of Science and Technology of China in Hefei, China from 1999 to 2004, and graduated with a Bachelor of Science degree in life sciences in 2004. He came to the University of Rochester in the fall of 2004 to pursue his graduate studies in the Brain and Cognitive Sciences Department. He received a Master of Arts degree in Brain and Cognitive Sciences in 2009 and a Doctor of Philosophy degree in Brain and Cognitive Sciences in 2010.

Diane Massam (Ph.D. MIT 1985) is Professor of Linguistics at the University of Toronto where she served as Chair of Linguistics from 2002 to 2008. Her research focus is on syntactic theory, in the areas of argument structure, case, predication, and word order, working mainly on Niuean, a Polynesian language. She has published papers in many journals such as Lingua, Oceanic Linguistics, Natural Language and Linguistic Theory, English Language and Linguistics, and Syntax, and has co-edited a book on ergativity with Springer. She has held several research grants from the Social Sciences and Humanities Research Council of Canada, and was co-editor of Squibs for Linguistic Inquiry (1998–2002), honorary research fellow at the University of Auckland (2001), visiting professor at Harvard University (2006), and an Erskine Fellow at the University of Canterbury, New Zealand (2010).

Eric Mathieu completed his Ph.D. at University College London in 2002. He is currently an Associate Professor in the Department of Linguistics, at the University of Ottawa. His research focuses on two languages: French and Ojibwe (an Algonquian language). He is a syntactician whose research has over the years concentrated
on the structure of the noun phrase, wh-movement and wh-in-situ, the mass/count
distinction, noun incorporation and denominal verb formation. His research is
published in journals such as Linguistic Inquiry, International Journal of American
Linguistics, Journal of Linguistics, Lingua, Probus and Studia Linguistica. He is the
editor of a special edition of Lingua on noun incorporation, and the author of
numerous chapters in books (with Oxford University Press, John Benjamins and
Springer, among others).

Ileana Paul is currently Associate Professor at the University of Western Ontario
(Canada) and she also holds the Canada Research Chair in Linguistics. Her research
focuses on the syntax of Malagasy, the language spoken in Madagascar. She has
published several book chapters and journal articles (in journals such as Syntax,
Lingua, and Oceania Linguistics) on a range of topics, including existentials, nominal
structure and interpretation, and clefts.

Francis Jeffry Pelletier’s Ph.D. was from UCLA in 1971, and he moved to the
University of Alberta that year. He was a Professor of both Philosophy and Com-
puting Science until 2004, when he moved to Simon Fraser University. At SFU he
was a Professor of both Philosophy and Linguistics, and a Canada Research Chair in
Cognitive Science. He moved back to the University of Alberta in 2009, where he is a
long-term visitor in the Philosophy department. He has authored or (co-)edited five
books (Mass Nouns: Some Philosophical Problems [1979]; Plato, Parmenides and the
Semantics of Not-Being [1990]; The Generic Book [1995]; Reference and Quantifica-
tion [2004]; Kinds, Things, and Stuff [2010]). He is currently working on a textbook
covering the history of logic.

Michael K. Tanenhaus is the Beverly Petterson Bishop and Charles W. Bishop
Professor of Brain and Cognitive Sciences and Linguistics at the University of
Rochester. Tanenhaus’ research focuses on real-time language processing, including
speech perception, spoken and visual word recognition, syntactic processing, refer-
ence resolution and generation, prosody, and conversation in collaborative tasks.
Some of his recent research is reviewed in: Tanenhaus, M.K. and Brown-Schmidt, S.
(2008), ‘Language processing in the natural world’, in Moore, B.C.M., Tyler, L.K. &
Marslen-Wilson, W.D. (eds.) The perception of speech: from sound to meaning,
Philosophical Transactions of the Royal Society B: Biological Sciences, 363, 1105–1122.

Heike Wiese (PhD 1997, Humboldt-University Berlin) is Professor of Contemporary
German Language and the Speaker of the Centre for Language, Variation, and
Migration at Potsdam University, Germany. Her research interests are in linguistic
variation and the syntax–semantics interface. In the domain of mass/count nouns,
her main work has been on nominal number marking from a typological perspec-
tive, on the grammatical mass/count distinction in language processing (with Maria
Piñango, Yale University), and on the conceptual foundations of number words (Numbers, Language, and the Human Mind, Cambridge University Press).

**Martina Wiltchko** is currently Associate Professor at the University of British Columbia, Canada. Her main research area is in syntax and how it interfaces with morphology and semantics. Her main research interest lies in the cross-linguistic variation of the realization of grammatical categories such as tense, number, person, and aspect. Languages she explores and does fieldwork on include Halkomelem (Salish), Blackfoot (Algonquian), and Ktunaxa (Kutenai, isolate).

**Niina Ning Zhang** is a Professor at the Linguistics Institute, National Chung Cheng University, Chiayi, Taiwan. Her major research interests include syntactic theories of generative grammar, morphology, the syntax-semantics interface, and the syntax of Mandarin Chinese. She has published works on the syntax of coordination and the syntax of numeral classifier constructions.
# Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>#</td>
<td>number</td>
</tr>
<tr>
<td>π</td>
<td>person</td>
</tr>
<tr>
<td>1</td>
<td>first person</td>
</tr>
<tr>
<td>2</td>
<td>second person</td>
</tr>
<tr>
<td>2/1</td>
<td>2nd person acting on 1st person</td>
</tr>
<tr>
<td>3</td>
<td>third person</td>
</tr>
<tr>
<td>ABL</td>
<td>ablative</td>
</tr>
<tr>
<td>ABST</td>
<td>abstract</td>
</tr>
<tr>
<td>ACC</td>
<td>accusative</td>
</tr>
<tr>
<td>ADV</td>
<td>adverb</td>
</tr>
<tr>
<td>AGGLOM</td>
<td>agglomeration</td>
</tr>
<tr>
<td>ANIM</td>
<td>animate</td>
</tr>
<tr>
<td>ARG</td>
<td>argument</td>
</tr>
<tr>
<td>ASP</td>
<td>aspect</td>
</tr>
<tr>
<td>AT</td>
<td>actor topic</td>
</tr>
<tr>
<td>AUX</td>
<td>auxiliary</td>
</tr>
<tr>
<td>CARD</td>
<td>cardinal</td>
</tr>
<tr>
<td>CARDP</td>
<td>cardinal phrase</td>
</tr>
<tr>
<td>CL</td>
<td>classifier</td>
</tr>
<tr>
<td>CLP</td>
<td>classifier phrase</td>
</tr>
<tr>
<td>CLT</td>
<td>clitic</td>
</tr>
<tr>
<td>CNP</td>
<td>common noun phrase</td>
</tr>
<tr>
<td>COL</td>
<td>collective</td>
</tr>
<tr>
<td>CONT</td>
<td>continuous</td>
</tr>
<tr>
<td>CONTN</td>
<td>continuative</td>
</tr>
<tr>
<td>COP</td>
<td>copula</td>
</tr>
<tr>
<td>CT</td>
<td>circumstantial topic</td>
</tr>
<tr>
<td>DAT</td>
<td>dative</td>
</tr>
<tr>
<td>DE</td>
<td>associative particle</td>
</tr>
<tr>
<td>DECL</td>
<td>declarative</td>
</tr>
<tr>
<td>DEF</td>
<td>definite (determiner)</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>--------------------------------------</td>
</tr>
<tr>
<td>DEM</td>
<td>demonstrative</td>
</tr>
<tr>
<td>DET</td>
<td>determiner</td>
</tr>
<tr>
<td>DIM</td>
<td>diminutive</td>
</tr>
<tr>
<td>DIR</td>
<td>direct theme marker</td>
</tr>
<tr>
<td>DP</td>
<td>determiner phrase</td>
</tr>
<tr>
<td>DUR</td>
<td>durative</td>
</tr>
<tr>
<td>ERG</td>
<td>ergative</td>
</tr>
<tr>
<td>EZ</td>
<td>ezafe</td>
</tr>
<tr>
<td>FEM</td>
<td>feminine</td>
</tr>
<tr>
<td>FOC</td>
<td>focus</td>
</tr>
<tr>
<td>FUT</td>
<td>future</td>
</tr>
<tr>
<td>GEN</td>
<td>genitive</td>
</tr>
<tr>
<td>HUM</td>
<td>human</td>
</tr>
<tr>
<td>IMP</td>
<td>imperative</td>
</tr>
<tr>
<td>IMPF</td>
<td>imperfect</td>
</tr>
<tr>
<td>IN</td>
<td>inanimate</td>
</tr>
<tr>
<td>INDEF</td>
<td>indefinite</td>
</tr>
<tr>
<td>INDEP</td>
<td>independent order</td>
</tr>
<tr>
<td>INF</td>
<td>infinitive</td>
</tr>
<tr>
<td>INST</td>
<td>instrumental</td>
</tr>
<tr>
<td>INT</td>
<td>intensifier</td>
</tr>
<tr>
<td>INVIS</td>
<td>invisible</td>
</tr>
<tr>
<td>LOC</td>
<td>locative</td>
</tr>
<tr>
<td>Q</td>
<td>quantifier</td>
</tr>
<tr>
<td>MASC</td>
<td>masculine</td>
</tr>
<tr>
<td>MS</td>
<td>milliseconds</td>
</tr>
<tr>
<td>n</td>
<td>light noun</td>
</tr>
<tr>
<td>N</td>
<td>noun</td>
</tr>
<tr>
<td>NEG</td>
<td>negation</td>
</tr>
<tr>
<td>NEUT</td>
<td>neuter</td>
</tr>
<tr>
<td>NHUM</td>
<td>nonhuman</td>
</tr>
<tr>
<td>NMZ</td>
<td>nominalizer</td>
</tr>
<tr>
<td>NOM</td>
<td>nominative</td>
</tr>
<tr>
<td>NP</td>
<td>noun phrase</td>
</tr>
<tr>
<td>NUMEP</td>
<td>numeral phrase</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>------------------------------------</td>
</tr>
<tr>
<td>O</td>
<td>object</td>
</tr>
<tr>
<td>OBV</td>
<td>obviative</td>
</tr>
<tr>
<td>OM</td>
<td>object marker</td>
</tr>
<tr>
<td>P</td>
<td>preposition</td>
</tr>
<tr>
<td>PART</td>
<td>participle</td>
</tr>
<tr>
<td>PASS</td>
<td>passive</td>
</tr>
<tr>
<td>p.c.</td>
<td>personal communication</td>
</tr>
<tr>
<td>PERF</td>
<td>perfect</td>
</tr>
<tr>
<td>PL</td>
<td>plural</td>
</tr>
<tr>
<td>POD</td>
<td>(phonetic) point of disambiguation</td>
</tr>
<tr>
<td>POSS</td>
<td>possessive</td>
</tr>
<tr>
<td>PP</td>
<td>preposition phrase</td>
</tr>
<tr>
<td>PRED</td>
<td>predicate</td>
</tr>
<tr>
<td>PREF</td>
<td>prefix</td>
</tr>
<tr>
<td>PRES</td>
<td>present</td>
</tr>
<tr>
<td>PFV</td>
<td>perfective</td>
</tr>
<tr>
<td>Pro</td>
<td>pronoun</td>
</tr>
<tr>
<td>PROG</td>
<td>progressive</td>
</tr>
<tr>
<td>PROX</td>
<td>proximate</td>
</tr>
<tr>
<td>PS</td>
<td>possessive</td>
</tr>
<tr>
<td>PST</td>
<td>past</td>
</tr>
<tr>
<td>Q</td>
<td>interrogative</td>
</tr>
<tr>
<td>QP</td>
<td>quantifier phrase</td>
</tr>
<tr>
<td>QU</td>
<td>quantifier</td>
</tr>
<tr>
<td>RECIP</td>
<td>reciprocal</td>
</tr>
<tr>
<td>REFL</td>
<td>reflexive</td>
</tr>
<tr>
<td>S</td>
<td>subject</td>
</tr>
<tr>
<td>SUBJ</td>
<td>subjunctive</td>
</tr>
<tr>
<td>SFP</td>
<td>sentence final particle</td>
</tr>
<tr>
<td>SG</td>
<td>singular</td>
</tr>
<tr>
<td>SUF</td>
<td>suffix</td>
</tr>
<tr>
<td>TMA</td>
<td>tense, mood, aspect marker</td>
</tr>
<tr>
<td>TOP</td>
<td>topic</td>
</tr>
<tr>
<td>TRANS</td>
<td>transitive</td>
</tr>
<tr>
<td>TT</td>
<td>theme topic</td>
</tr>
</tbody>
</table>
### Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>v</td>
<td>light verb</td>
</tr>
<tr>
<td>V</td>
<td>verb</td>
</tr>
<tr>
<td>VP</td>
<td>verb phrase</td>
</tr>
<tr>
<td>VTA</td>
<td>transitive animate verb</td>
</tr>
<tr>
<td>VTI</td>
<td>transitive inanimate verb</td>
</tr>
</tbody>
</table>
1

The count mass distinction: Issues and perspectives

JILA GHOMESHI AND DIANE MASSAM

1.1 Setting the stage

This volume explores the expression of the concepts count and mass in human language. It not only serves to expand our understanding of the key issues raised by these concepts, but also significantly contributes to their cross-linguistic exploration, as it includes papers with a focus on Blackfoot, Cantonese, Dagaare, English, Halkomelem, Lithuanian, Malagasy, Mandarin, Ojibwe, and Persian, as well as discussion of several other languages including Armenian, Hungarian, and Korean. The linguistic subfields covered in the papers are also diverse, including syntax, morphology, semantics, acquisition, and processing. In this introduction we will discuss the concepts of count and mass and their role in grammar, presenting the range of issues and points of view both within the existing literature and within the papers in this volume. Exhaustive coverage of the literature is not attempted in this introduction, however an extensive bibliography is provided by the sum of this volume.

The concepts count and mass are far from straightforward but in order to begin our consideration, let us develop a relatively simple working definition. The concepts are first and foremost relevant in the nominal domain, so we start there, and consider a count noun to be one that identifies a unit that can be counted, such as dog or table (one dog, three tables), and a mass noun to be one that names an entity that comes in mass form and therefore cannot inherently be separated into countable units, at least not without a change in meaning (#one air, #six rices). Along the same lines, two criteria, divisibility (Cheng 1973) and cumulativity (Quine 1960) can be invoked (Gillon 1992). Mass noun referents can generally be divided without loss of integrity, that is, half an amount of water still yields water, which is certainly not the case for a dog; and mass nouns can also be accumulated without essential change. A further test in English is that count nouns are those that can be pluralized
(dogs), whereas mass nouns are those that cannot be pluralized (without a change in meaning) (#rices). Once we have established this basic (partly English-based) distinction, we can pose three core questions: (1) Do these concepts form part of a universally shared cognitive capacity and are they mapped uniformly to the real world? In other words, do all humans understand dogs to be essentially different from water, and if so, would all humans consider rice to be mass and pea to be count? (2) Are the concepts count and mass in some way grammatically encoded in all languages, or is it possible for languages to forego any grammatical expression of these concepts? (3) Are count and mass always realized in the same way in grammars cross-linguistically, such as through the availability (or not) of plurality noted for English above, and if not, what are the various ways that they are encoded in language?

Almost all research in the area of count and mass addresses one or more of these three questions. Because the count mass distinction encompasses grammatical, semantic, and conceptual categorization, it has key relevance for our understanding of the relationship between syntax, semantics, and the real world, and its study has therefore engaged morpho-syntacticians, semanticists, and philosophers, and these points of view are reflected in the papers in this volume. The count mass distinction is in fact a perfect laboratory for the exploration of the surprisingly complex relation between what at first glance seem to be universal incontrovertible aspects of meaning, and their extremely varied grammatical realizations.

In current theoretical literature, there are two main overall positions on the organization of the role of count and mass in human grammar. The first view, represented famously by Chierchia (1998b), is a typological approach. Chierchia (1998a, 1998b, cf. Krifka 1989) suggests that in some languages, all nouns are in some sense mass nouns (e.g. Chinese, Japanese), while in others, some nouns or noun extensions are count whereas others are mass (e.g. Romance, English). Although Chierchia states that there are no languages in which all nouns are count (see also Rothstein 2010), others have suggested that this might be possible, such as in Hopi (Whorf 1941), and Lillooet Salish (Davis and Matthewson 1999). In contrast to the typological approach, others have taken a universalist approach. Borer (2005) holds that cross-linguistically lexical items have very few inherent properties, and certain aspects of their interpretation depend entirely on context. In this view, a noun like dog, for example, becomes mass simply by virtue of appearing in the frame associated with mass nouns (in English, as a bare noun), and a noun like air becomes count by appearing in a count environment (in English, pluralized). Gillon (1992, 1999) also argues that the count mass distinction is grammatical, and that it cannot be based on real world properties of objects, presenting the arbitrary nature of pairs such as leaves and foliage. This purely syntactic approach has the virtue of explaining the Universal Grinder phenomenon whereby it is possible to say There was dog all over the floor (Pelletier 1975, cf. Bunt 1985a, 1979), but some might argue that it predicts
more coercibility than is actually found in language. A third position, taken by Pelletier (this volume), takes this potential coercibility to its extreme by proposing that all nouns are lexically both count and mass, thus resisting the kind of abstraction inherent in the two other positions. The nature of the interplay between the deep properties of the lexical items and their referents, and the syntactic frames they appear in is of central importance and feeds general questions in linguistic theory regarding the relation between the lexicon and syntax.

With this general background in place, let us discuss each of the three questions presented above.

1.2 Are count and mass conceptually universal and are they mapped to the real world uniformly?

Most scholars agree that the conceptual or ontological ability to make a distinction between count and mass is a universal one, which brings it in line with other semantico-grammatical categories such as gender and animacy that correspond in some way to real world distinctions. But if the categorization was inherently tied to the deep conceptual meaning of individual lexical items, we would expect to find language-internal and cross-linguistic uniformity in the matching of lexical items to the categories, which does not appear to be the case, as evidenced by Gillon’s (2009) many pairs, such as shoes and footwear. This situation leads us to wonder exactly which essential concepts are implicated by the count mass distinction. For example, the lack of isomorphy of class between item and expression might indicate that while there is a universal distinction, it might be other than count versus mass. The relevant distinction might instead be count versus non-count, allowing for collectives such as footwear and furniture (Wiese this volume, Rothstein 2010). This allows for a variety of types of non-count nouns, such as those with no minimal parts (e.g. water), and those collectives that can be divided, but behave as mass (e.g. cattle, furniture, spaghetti). However, Barner and Snedeker (2005, 2006) show that collectives pattern with plural count nouns conceptually. In answer to the question ‘Who has more...?’, respondents say five small forks are more than one large fork for category silverware, similarly to the count noun shoes, but five small mounds is not more than one huge mound for the category toothpaste. Such problematic mismatches of concept, expression, and behaviour might indicate that we still do not have the right ontology, or perhaps that we are not viewing it correctly (Allan 1980). Grimm (this volume) demonstrates with data from Dagaare that the count mass distinction is better understood as scalar rather than binary, with water on one end, ribs in the middle, and dog at the other end. Similarly, Zhang (this volume) argues that count and mass are not fine-grained enough categories to account for Mandarin Chinese, and further, that they do not constitute a binary opposition. She proposes
that two independent features, numerability and dimensionality, are at play to determine noun class membership across languages.

On the other hand, problematic mismatch cases might indicate simply that there is a high degree of grammatical mediation in the expression of the concepts, as we know there to be in other cases, such as gender and animacy, allowing similarly for language-internal and cross-linguistic variation. Wiltschko (this volume) considers this to be the case, arguing that count and mass are instantiations of the broader concept of nominal aspect, which might be expressed in a number of ways (e.g. by animacy), or not expressed at all, at the grammatical level. Cowper and Hall (this volume) also argue for the strong effect of grammatical mediation: through variations in featural syntax and morphology, languages can yield quite different effects. An additional muddying effect is that it is usually possible to coerce nouns into their opposite categories – two waters is fine in a restaurant order (Wiese and Maling 2005) and as noted above, count nouns can be put through the Universal Grinder (Pelletier 1975) to yield a mass reading, as discussed further by Cheng (this volume), Cheng, Doetjes, and Sybesma (2008), and Pelletier (this volume). Several scholars argue that the precise interpretation of coerced nouns is due to the mediating effect of grammar or the lexicon on count and mass (e.g. Wiese this volume, Borer 2005, Pelletier this volume).

1.3 Are count and mass universally expressed in language and are they always expressed in the same way?

As outlined briefly above, in considering how the concepts are expressed across languages, several scholars have relied on deep cross-linguistic variation, and taken the position that in some languages all nouns are primarily understood as mass (Chierchia 1998b), while in other languages, such as English, individual nouns can vary in this regard. This core difference is argued to result in surface distinctions such as the use of number or classificatory systems. Others have allowed for grammar to mediate in other ways, through lexical or syntactic specification. Taking a grammatical view of the count mass distinction entails the possibility for languages not to express it at all, but there has been relatively little exploration of this possibility. Wiltschko (2005b, this volume) fills this lacuna in her discussion of Halkomelem Salish, for which language she argues that the dichotomy simply does not exist. This forms part of her larger view, according to which languages can map a range of concepts, such as tense or number, to different places in the grammar, or they can not map them at all. Zhang (this volume) argues that there is not precisely a count mass distinction in Chinese, since all nouns are non-count, but that there is a somewhat different distinction of mass versus non-mass, evidenced through the classifier system, which utilizes morphemes to encode countable units of individuated items.
It is generally accepted that even in languages where there is a grammatical count mass distinction, there is variation in the means of expression of the distinction. This is made especially clear through the discussion of classifier versus number languages, touched on by virtually all who work on the issue of count and mass. Wiese (this volume) approaches this issue through the study of collectives, and argues that the availability of plural marking is not always a reliable diagnostic for a count mass distinction. The chapters in this volume contribute substantially to the growing evidence that there are a variety of ways to express a count mass distinction in language.

1.4 If count and mass are not expressed in the same way universally, how are they expressed?

Turning now to linguistic diagnostics of the count mass distinction, there are several grammatical phenomena to discuss. For example, in English, mass nouns appear bare, without determiners (*I like rice versus *I like table) and also cannot generally be pluralized or counted, whereas count nouns must appear with determiners (*I saw hat) and can be pluralized and counted (two hats). The relation between count mass and determiners is complex and much studied, stemming in recent work from Chierchia (1998a, 1998b), who makes a strong connection between the use of determiners in a language and its count mass system. There has been much exploration of this topic in Germanic and Romance languages, which differ in that Germanic languages do, and Romance languages do not generally permit bare plural mass nouns (e.g. Acquaviva 2008, Carlson 1977, Dayal 2004, Diesing 1992, Heycock and Zamparelli 2003, Krifka 1995, Longobardi 1994, 2001, Zamparelli 2002, among many others). In Persian (Ghameshi 2003, Ghanibadi this volume), the link between the determiner system and the count mass distinction is made via number marking in that plural marking on mass nouns is used to portion out, but is syncratic with definiteness. Klein et al. (this volume) provide experimental evidence for the link between the count mass system and the determiner system. They argue based on comprehension experiments that classifiers can function as gender marked determiners in their impact of referential selection, that is on the choice of the following noun, and they furthermore show an asymmetry between massifiers (denoting units of a mass) and classifiers in this regard.

As in Persian, in Ojibwe (Mathieu this volume), and Niuean (Massam 2009), nouns corresponding to English mass nouns can also be pluralized. It might be considered then that in such languages there is no count mass distinction, since the pluralization diagnostic does not appear to be sensitive to it. Mathieu (this volume) argues, however, that it is not the occurrence of plurality on putative mass nouns that matters, but rather, the interpretation of this plurality as a divider of stuff or as
filling some other function, such as giving a unit of measure reading with an atomization effect. In Ojibwe, he argues, plural attaches to a noun that has already been made singular, evidenced through a morphologically marked gender shift. Thus, he considers that different behaviours of plural morphemes may indicate differences in the meaning of plurality across languages, rather than differences in count and mass. Wiltschko (this volume) also addresses this question, arguing that in Halkomelem (Salish) and Blackfoot (Algonquian) number is more properly termed visibility or animacy.

Another way in which pluralization bears on the issue of the count mass distinction is in languages that lack number marking altogether on nouns. Paul (this volume) shows that Malagasy nouns exhibit general number, meaning that bare nouns can be construed as singular or plural and are not inflected for number. Nonetheless, Paul argues that there is a robust count mass distinction in Malagasy, based on quantifier distribution. She argues that the lack of functional structure cannot be construed the same way across languages.

In many languages in which number does not play a strongly grammatical role it is not the key decider of count and mass. In Bantu languages, count and mass nouns are marked with different noun class markers (Cheng and Schadeberg 2007). In Mandarin and other Chinese languages, distinctions between count and mass are made by means of a system of classifiers. In order to count a noun in Chinese, classifiers must be employed (e.g. two unit dog, three volume book). Classifiers thus individuate count nouns, but there are also massifiers in these languages, which form unindividuated groups (Cheng and Sybesma 1998, 2005, Cheng this volume, Zhang this volume, Klein et al. this volume), thus raising new questions about the count mass division.

The distinction between languages that use classifiers in counting expressions, such as Mandarin and Thai, versus languages that do not, such as English and French has been considered to constitute a basic cross-linguistic typological difference, relating as well to the ability for nouns to appear bare (Chierchia 1998b, Paul this volume, Piriyawiboon 2010). Bale and Barner (this volume) argue however, from a learnability point of view, that it is not the mere presence of morphemes such as classifiers or plurals that is crucial to differentiating language types but rather their relation to particular semantic interpretations (see also Mathieu this volume).

An alternate conclusion that is drawn from the exploration of typological properties of classifier languages is that classifiers and plurality are two sides of the same functional coin since both are used to provide division of stuff (Borer 2005, Mathieu this volume, Piriyawiboon 2010, Cowper and Hall this volume). With Borer (1985), Cowper and Hall argue that individuation is a cross-linguistic process, with number and classification systems being two ways to realize individuation, although they consider that the two can co-exist in a single language, as long as there is only one in a given nominal expression. The role of individuation is also explored by Grimm
(this volume). He considers individuation to be a scalar rather than a binary opposition, showing that number can be realized differently for different nouns, depending on their place on this individuation scale. Although for some, number and classifier systems are in complementary distribution, both expressing individuation, Cheng (this volume) argues that classifiers can differ cross-linguistically as to whether they play the role of a unit marker, comparing Mandarin versus Cantonese, and she also argues that classifiers can express number. Along similar lines, Massam (2009) examines Niuean, a language with both number and classifier systems, and suggests that in Niuean we see a language in the complex process of shifting from a classifier to a count system.

A final major issue taking us to the edges of the nominal delimitation of the count mass distinction is the relation between nominal number and sentential aspect (Rijkhoff 1991, Rothstein 2004, Schwarzschild to appear). Armoskaite (this volume) explores this issue through Lithuanian aspectual prefixes, showing that number is not restricted to the nominal domain but plays a role in the verbal domain as well. Borer (2005), in her discussion of bare plurals, similarly explores this issue, paving the way for a unified description of quantity within the nominal and the aspectual domain, through Krifka’s (1989) notion of quantization.

1.5 Conclusion

Let us end by summing up our answers to the three questions posed at the outset. For the first one, the overall consensus expressed in this volume is that while the general concepts of count and mass are available to all humans, lexicalization or grammaticalization forms a key role in whether a given nominal is treated as a count or mass noun. There is also evidence to suggest that in some languages, the concepts receive no grammatical realization, possibly as expected given that other core concepts like gender and plurality appear to be similarly variable in their grammaticalization cross-linguistically. The lack of a grammatical distinction may also reflect the fact that count vs. mass is just one possible realization of a deeper and broader concept, perhaps related to the categories of nominal and verbal aspect. And finally, it is very clear that count and mass are not always expressed in the same way cross-linguistically, even if they are present in a given grammar, so that studying this dichotomy provides insight into other notions such as number and determination as well. The papers in this volume bring us significantly further down the road in our understanding of the various facets of cross linguistic variation in the count mass domain. At the same time, we also see haunting similarities across distantly related groups of languages, which suggests that there might be underlying unity between the varying forms used to delineate the count mass distinction.

Before concluding, we would like to pay tribute to the inception of this volume, which grew out of the Workshop on Count and Mass Nouns held at the University
of Toronto in February 2009, organized by Nattaya Piriyawiboon and Diane Massam. Funding for the workshop was provided by the Social Sciences and Humanities Research Council of Canada. There were several presentations at the workshop that could not be submitted for this volume but they deserve credit here as they influenced the papers that do appear. These were presentations by Hagit Borer, Gennaro Chierchia, Brendan Gillon, So-One Hwang, Kimiko Nakanishi and Elizabeth Ritter, and George Tsoulas and Eytan Zweig. Other acknowledgements must also be made. The volume would not have seen the light of day without the remarkable contributions of several people. We would like to thank Nattaya Piriyawiboon in particular, as she worked hard and successfully to secure funding for the Workshop on Count and Mass Nouns and has influenced this volume in many ways. We would also like to thank the many students who volunteered to help with the workshop, and David Adger, Hagit Borer, Jan Chamier, Gennaro Chierchia, Sarah Clarke, Anna Maria Di Sciullo, Victoria Hart, Michela Ippolito, Giuseppe Longobardi, Jennifer Lunsford, Yves Roberge, Anne St-Amand, Julia Steer, the Oxford University Press, and the Department of Linguistics at the University of Toronto. The workshop was the fifth in an ongoing series of Canadian workshops on issues related to nominal phrases, and we would also like to thank the core participants in these events for their continuing collegiality, encouragement, and inspiration.
Lexical nouns are both +MASS and +COUNT, but they are neither +MASS nor +COUNT

FRANCIS JEFFRY PELLETIER

2.1 Introduction: Informal accounts of +MASS and +COUNT

Many of the writings on the mass/count distinction, whether they are from the ESL literature, the descriptive grammar literature, or the more theoretical linguistics literature, explain this distinction by mentioning some paradigm examples. It will be said that the lexical words:

(1) [+MASS] water, blood, air, sand...

are +MASS, whereas:

(2) [+COUNT] person, dog, tree, house...

are +COUNT. Sometimes, although not always, it is also pointed out that some ‘abstract’ nouns are +mass and others are +count:

(3) [+ABST, +MASS] advice, knowledge, curiosity, software...

(4) [+ABST, +COUNT] suggestion, belief, apology, program...

And sometimes it is mentioned that complex noun phrases are also to be categorized mass/count:

(5) [+MASS] dirty water, red blood that is on the floor, too much justification...

(6) [+COUNT] tall person, big dog that is sleeping in the corner, each house on the street...
although the majority of the literature discusses mass/count only as it occurs with lexical nouns.¹

Works on the distinction as it occurs in English typically add a few syntactic conditions that distinguish \( + \text{mass} \) from \( + \text{count} \):

(7)  
\begin{enumerate}
\item Count nouns, but not mass nouns, have plural forms and thus can agree with plural verbs.
\item Count nouns, but not mass nouns, can occur with numerals and counting phrases.
\item Singular count nouns, but not mass nouns, employ the quantifiers each, every, (stressed quantifier) some, and indefinite \( a(n) \).
\item Plural count nouns, but not mass nouns, employ the quantifiers few, several, many.
\end{enumerate}

(8)  
\begin{enumerate}
\item Mass nouns, but not count nouns, do not have plural forms and thus all verb agreement is singular.
\item Mass nouns, but not singular count nouns, can occur with measure phrases like liters of, amount of.
\item Mass nouns, but not count nouns, employ the quantifiers much, little.
\item Mass nouns, but not singular count nouns, employ the unstressed some and the quantifier most.
\end{enumerate}

And finally, there are a number of semantic features about \( + \text{mass} \) vs. \( + \text{count} \) nouns that should hold in any language. They should hold cross-linguistically because these features are taken to characterize either the reality that the nouns denote (externalism) or the mental item that is occasioned by the term (internalism), depending on the semantic theory under consideration. In either case they not seen as describing aspects of a language. For example,

(9)  
\begin{enumerate}
\item Count nouns, but not mass nouns, designate a set of (countable) entities.
\item Mass nouns, but not count nouns, designate stuff.
\end{enumerate}

There are two basic approaches to theories of \( + \text{mass} \) nouns: one takes the syntactic criteria as basic while the other takes the semantic criteria as basic, in their way of making the mass/count distinction. Of course, each thinks that the other criterion will naturally follow from the one they choose as basic. But as we will see, this just does not happen.²

¹ There is also a literature on extending the distinction to lexical verbs and to verb phrases, but we will not follow that up here. See works on ‘eventology’, e.g. Mourelatos (1978), Bach (1986a,b), for this extension.

² There are also some who think these two rationales for calling some lexical noun \( + \text{mass} \) or \( + \text{count} \) somehow reach a détente. For instance, a theory might declare the semantic criteria ‘really’ to be syntactic,
2.2 \( +\text{MASS} \) and \( +\text{COUNT} \) as syntax

Many descriptive grammars of English, e.g. Quirk et al. (1985), give a syntactic characterization of the mass/count distinction within the category of noun. That is, they view the fact that some noun (e.g. water) is a mass term as giving an explanation for why some combinations with other words are ungrammatical. For example, they might put forward the criteria in (7) and (8) as the rationale for classifying the paradigm examples cited above in (1–6).

The \( +\text{COUNT} / +\text{MASS} \) features are viewed by Quirk et al. and others of this syntactic persuasion to be a part of the lexical characterization of the nouns. These features are to be inherited from the lexical items into the larger and larger syntactic units that are present in extended phrases. So, blood as a lexical entry contains the syntactic feature \( +\text{MASS} \), and this is inherited by the common noun phrases (CNPs) bright red blood and bright red blood that is on the floor and the full determined phrase the bright red blood that is on the floor. The fact that this longer phrase is also \( +\text{MASS} \) is what ultimately explains why

\[
\begin{align*}
(10) & \quad \text{a. *The bright red blood that is on the floor are slippery.} \\
& \quad \text{b. *Each bright red blood that is on the floor is slippery.}
\end{align*}
\]

are ungrammatical. (10a) is ungrammatical because the subject phrase being \( +\text{MASS} \) prohibits it from being plural, as (8a) says, and hence the agreement with the verb phrase fails. (10b) is ungrammatical because the determiner each is not allowed to combine with the \( +\text{MASS} \) common noun phrase bright red blood, as (7c) says.) Violations of the constraints involving \( +\text{MASS} \) and \( +\text{COUNT} \) yield ungrammatical results that have the same status as other syntactic violations; (10a) and (10b) are no more a part of English than are

\[
\begin{align*}
(11) & \quad \text{a. *Person now rash.} \\
& \quad \text{b. *That fact the because.}
\end{align*}
\]

2.3 \( +\text{MASS} \) and \( +\text{COUNT} \) as semantics

Some descriptive grammars of English, e.g. Huddleston and Pullum (2002), think of the mass/count distinction as a description of the semantic properties of the denotation of the terms. In this type of view, mass meanings contrast with count meanings because of (9). In turn, this general semantic categorization has some more particular manifestations in the meanings of lexical items:

and that some nouns will be treated as (syntactically) \( +\text{MASS} \) or \( +\text{COUNT} \) on account of the original syntactic reasons while others will be treated as (syntactically) \( +\text{MASS} \) or \( +\text{COUNT} \) on account of these new reasons. See Wiltshire (this volume) for treating \( \pm\text{BOUNDED} \) in this manner.
(12) a. Mass meanings are *divisive in their reference*; count meanings are *true of a unit as a whole*.
   b. Mass meanings are *cumulative in their reference*; (singular) count meanings are *not true of groups of that which they are true of*.
   c. Stuff that mass meanings are true of *cannot be counted*; count meanings are *true of individuated items that can be counted*.
   d. Stuff that mass meanings are true of *can be measured*; (singular) count meanings are *not measurable*.

Some theorists take the divisiveness and the cumulativity conditions together to be called the *homogeneous in reference* condition.

The fundamental difference between mass and count terms is that count terms are true of *objects*—entities that are distinct from each other even while being of the same type, and thus one can distinguish and count them—while mass terms are true of *stuff* that is undifferentiated with respect to the term being used to describe it. This in turn explains why mass terms, unlike count terms, are *divisive* in their reference: they permit something that the mass term is true of to be arbitrarily subdivided and the term to be true of these parts as well. Taking the water in the glass to be something that *is water* is true of, it can be divided into parts and *is water* will be true of both parts. And again, mass terms, unlike count terms, are also *cumulative* in their reference: putting the water contained in two glasses into a bowl yields something of which *is water* is true. But the same is not the case with a count term like *dog*. Chopping up a dog does not yield more things of which *is a dog* is true, nor do two dogs make a thing of which *is a dog* is true.3

In a semantic approach, the features \(+\text{mass}/+\text{count}\) are descriptions of the semantic value of lexical nouns and the larger CNPs and the still larger determiner phrases (DPs) and noun phrases (NPs), etc. Thus, they do not figure in the syntactic well-formedness constraints of a grammar, but would emerge as a description of what the semantic values of the embedded nouns are, and how these semantic values get altered by the syntactic combination of those nouns with other words. In such a picture, these features do not syntactically rule anything out; the most that can be said is that certain combinations are ‘semantically anomalous’, and hence can’t be interpreted. According to the semantic theory, the violations alleged by the syntactic theory are ‘really’ a consequence of the fact that ‘it wouldn’t make any sense’ to try to interpret the violations. For example, trying to interpret ‘each mud’ is impossible because ‘each’ requires a domain of entities but ‘mud’ doesn’t provide such a domain.

3 Other than in a Frankenstein-like scenario.
As with the syntactic version of $+$mass$/+$count, in the present semantic theories, the lexicon supplies individual words with a set of syntactic features and also a set of semantic values. Larger and larger phrases that contain the noun also contain the semantic information mentioned in the lexical items, modified in accordance with rules that describe the semantic effect of being syntactically combined in the manner that is employed. In this type of theory, sentences that violate the ‘appropriateness’ of the semantic features of $+$mass and $+$count are seen as grammatical but not interpretable. So the sentence (10b) would not be ungrammatical for using an ‘individuative quantifier’ with a mass noun phrase—it would only be ‘uninterpretable.’ (10a), on the other hand, would be marked as syntactically ill-formed, this time on the grounds that the singular subject has a number mismatch with the verb phrase.

The difference between $+$mass$/+$count as syntax and $+$mass$/+$count as semantics thus is whether these features are seen as syntactic well-formedness constraints that yield ungrammaticality when violated or as semantic interpretability constraints upon syntactically correct sentences that yield semantic anomaly when violated.

2.4 Problems with the syntactic approach

The syntactic approach is, well, syntax-driven. The lexical items are assigned either a $+$mass or $+$count feature, and this feature controls the syntactic admissibility or inadmissibility of larger phrases. But there are many words that have both mass and count meanings, for instance:

(13) a. Concrete terms
    (i) a lot of chocolate / many more chocolates
    (ii) more discipline / an academic discipline
    (iii) too much paper / write a paper
    (iv) drink beer / drink a beer

b. Abstract terms
    (i) much discussion / three different discussions
    (ii) much justification / many justifications
    (iii) a lot of difference / two differences
    (iv) much more data / many more data

The examples in (13) are just the tip of the iceberg. There are many more of the ‘dual life’ terms which have been illustrated in (13) above, sometimes forming regular patterns, but sometimes not:
(14) Mass terms used ‘countily’:
   a. Pinot Noir is wine / Pinot Noir is a wine
   b. Kim produces sculpture / Kim is producing a sculpture
   c. Sandy likes lamb / Sandy likes every lamb
   d. Beer on the table / Three beers on the table / Eight beers on tap

(15) Count terms used ‘massily’:
   a. Leslie has more car than garage
   b. Chis Pronger, 6’6” worth of ice-hockey defenseman . . .
   c. He’s got woman on his mind
   d. What a hunk of man!
   e. Some people like data better than theory

As Huddleston and Pullum (2002: p.335) remark ‘the dual use of chocolate is not remotely exceptional but is representative of an extremely widespread phenomenon’, and they follow this with a list of 25 examples chosen over a wide variety of types of nouns that illustrate just how wide-spread is the phenomenon of a noun having two equally-salient meanings where one is +mass but the other +count.

And then there’s the ‘universal grinder’ of Pelletier (1975), which is like a meat grinder except that it can accommodate any object, no matter how large, and its teeth are so powerful and fine that it can grind anything, no matter how strong. One inserts an object that falls under any count noun into one side . . . for example, a hat. Push the button, and the result is that there is hat all over the floor.4 Another push of the button and we can have book all over the floor. An unfortunate accident might generate curious cat all over the floor.

One might also think of ‘universal packagers’ in this regard, that take any item of which a mass term is true and convert it into an object. Any time there is a use for a particular type of some mass then there can be a count term that describes it—for example, *a finely-silted mud*, which can be a name for a type of mud and also a predicate that is true of all individual exemplars of this type. And if there is a standardized amount of *M* that is employed in some use, then there will be a count term that describes this amount, such as *a beer* or *an ice cream*. Furthermore, there seems always to be a count use for any alleged mass term *M*, meaning (roughly) *a kind of M*. Putting all these together, a term like *a scotch* could be true of individual servings (thus being independently true of each piece of the actual matter in the various glasses), or true of the different standardized amounts (so that two instances of the same standard one-ounce serving count as only one such standardized amount), or true of the different kinds of scotch on the table or available at the bar. Thus any of ‘one’, ‘three’, ‘five’ could be true answers to the question ‘How many

---

4 This is true despite the fact that we might have some other term, e.g. *felt*, that also describes what is on the floor.
scotches are on the table?’, when the table contains five glasses poured from three
different kinds of scotch, each glass containing the standard double shot.

These considerations show that the appropriate theory needs to talk about
meanings of terms, or uses of the terms, or maybe occurrences thereof (some
occurrences are $+$mass, others of the same word are $+$count). But then this is
no longer a syntactic account. And the syntactic approach just doesn’t work. For, it
will turn out that since any noun can be either mass or count, a $+$mass/$+$count
syntactic distinction does not work; nothing is ruled out purely by the syntactic rules
involving $+$mass and $+$count.

2.5 Problems with the semantic approach

As we have seen above, many or most or almost all nouns have both a natural mass
and a natural count sense. So the basic lexical item that gets entered into a phrase
structure description of a sentence will be one of these senses. It is never very clear
how this is supposed to be effected in a grammar, but we will not pause over that
here, and simply assume that there is some way that this can be done. But even if we
can assume this, there nonetheless seem to be some serious difficulties that are
semantic mirrors of the difficulties found in the syntactic approach.

Many formal semanticists (e.g. Link 1983, Chierchia 1998a, b, Pelletier and
Schubert 1989/2003) take the characteristics in (12) to be best accounted for in
terms of a semi-lattice theory. A semi-lattice has no lowest elements and is atomless.
The idea is that anything that water might be true of has subparts—things in the
lattice that are its parts—of which water is also true; and any two elements in the
water-lattice find a joined element also in the lattice that represents the merge of
those two elements.

But it should be noted that many mass terms obviously are not atomless in the
sense required by this theory. Consider:

\[(16) \text{furniture, cutlery, clothing, equipment, jewelry, crockery, silverware, footwear,}
\text{bedding, toast, stemware, gravel…}\]

Clearly there are atomic parts of these, and yet they are considered mass terms by
any of the traditional grammars. So it cannot be an atomless mereology that
accounts for the mass nature of these words; and by extension, since it doesn’t
account for the mass nature of these particular words, there seems to be no reason to
think it is responsible for the mass nature of any words.

Some theorists, e.g. Huddleston and Pullum (2002), take this as evidence that
terms like those in (16) are of a different nature than what we have been calling ‘mass
terms’, and are to be treated differently. Huddleston and Pullum call them ‘aggregate
terms’ and semantically distinguish them from other mass terms by virtue of their
being true of ‘very different sorts of things’. The idea is that furniture, for example, is
true of sofas, chairs, tables, carpets, and so on, and that these are ‘very different’ from one another. But a true mass term, for example blood, is really true only of one kind of thing.

But one might still wonder: are there any words at all that obey the condition on divisiveness? Or put another way, are there really any words that are atomless—whose referent has no smallest parts? Doesn’t water, for example, have smallest parts: H$_2$O molecules perhaps? Certainly coffee and blood have smallest parts, as do other mixtures. A standard defense of the divisiveness condition in the face of these facts is to distinguish between ‘empirical facts’ and ‘facts of language’. It is an empirical fact that water has smallest parts, it is said, but English does not recognize this in its semantics: the word water presupposes infinite divisibility.

It is not clear that this is true, but if it is, the viewpoint suggests interesting questions about the notion of semantics. If water is divisive but water isn’t, then water can’t be the semantic value of water (can it?). In turn this suggests a notion of semantics that is divorced from ‘the world’, and so semantics would not be a theory of the relation between language and the world. But it also would seem not to be a relation between language and what a speaker’s mental understanding is, since pretty much everyone nowadays believes that water has smallest parts. Thus, the mental construct that in some way corresponds to the word water can’t be the meaning of water either. This illustrates a kind of tension within ‘natural language metaphysics’.

Further problems with the semantic approach to the mass-count distinction come from the fact that there are pairs of words where one is mass and the other is count and yet the items in the world that they describe—or in the minds of speakers using the terms—seem to have no obvious difference that would account for this. On the intuitive level, it seems that postulating a semantic difference should have some reflection in the items of reality that the terms designate (or in the mental life of speakers using the terms). But this is just not true. There seems to be nothing in the referent (or speaker-beliefs/intentions) of the following mass vs. count terms that would explain how they should be distinguished, as they intuitively are (see McCawley 1975 for further examples).

(17) a. Concrete terms
   (i) baklava vs. brownies
   (ii) spaghetti vs. noodles
   (iii) garlic vs. onions
   (iv) rice vs. beans

5 At least, there are volumes that contain coffee, and there are subvolumes of such a volume which are so small that they do not contain coffee. And so some sort of ‘continuity principle’ suggests that there is a cut-off line or interval that yields smallest parts of coffee.

6 For a description of, and defense of approaching metaphysics this way, see Bach (1986a,b).
b. Abstract terms
  (i) success vs. failures
  (ii) knowledge vs. beliefs
  (iii) flu vs. colds

To many, these examples and their surrounding facts have seemed to prove that the linguistic features of +count and +mass do not have any backing in reality. Nor any backing in people’s intuitive understanding of when a word will be +mass or +count nor what it is for a word to be +mass or +count.

2.6 Evaluation, and a flaw in common

Sections 2.4 and 2.5 have presented a litany of problems with the two general approaches. I wish to point the direction to the kind of theory that can avoid the listed difficulties. In particular, I want to avoid the syntactic approach’s use of syntactic features that don’t ever make any construction be ungrammatical. And I want to avoid the semantic approach’s view that the (alleged) violations are not syntactic claims, for example its claim that three water is just ‘semantically anomalous’ and not syntactically ill-formed. I also would like to challenge the semantic approach’s claim that there is some deep ontological backing to the distinction between +mass and +count, and challenge the internalist view that there is some deep conceptual backing to the distinction.

What the syntactic and semantic views have in common is that they make +count/+mass be features of lexical nouns. It could instead be a ‘constructional feature’ introduced when CNPs are formed into DPs and NPs. My goal is to have syntactic features that give rise to ill-formedness when violated, and to have semantic consequences follow from these constructions. These semantic consequences could be described in terms of some semantic features that somehow mirror the syntactic ones.

2.7 A different approach

In this section I’ll lay out an approach that honors what I think is right about the syntactic approach and what I think is right about the semantic approach, but which avoids the flaws of both. The overall view is that lexical nouns are neither mass nor count, but that they are both mass and count. The next two subsections explain how this can happen.

2.7.1 Lexical nouns are neither +mass nor +count

Syntactically speaking, the proposal is that lexical nouns do not have any syntactic marker for +mass or +count, but the phrases they occur in can become marked
either +mass or +count. An example, simplified and not containing any other factors than mass/count, goes like this:7

(i) beer lexically lacks any syntactic feature of +mass/+count.
(ii) dark beer, beer on the table (CNPs) lack any syntactic feature of +mass/+count.
(iii) beers (a CNP) has the syntactic feature +count.
(iv) is beer (PRED) has the syntactic feature +mass.
(v) a beer, many beers (NPs) have the syntactic feature +count.
(vi) a lot of beer, beer (DPs/NPs) have the syntactic feature +mass.

+mass and +count are not lexical features: no lexical item has them. An alternative way of putting this is that lexical nouns are unspecified for the syntactic features +mass/+count. In this way the existence of ‘dual life’ nouns is no problem—there aren’t any, really. Instead, sometimes the noun syntactically composes to form a larger phrase that is +mass, sometimes it forms one that is +count. Chocolate, for example, is neither +mass nor +count, but composed with a to form a chocolate, the resulting NP is marked as +count; when composed with much to form much chocolate, the resulting NP is marked +mass.

As well, the observed syntactic violations that are attributed to violations of the restrictions on +mass and +count are honored: they do in fact occur, and the violations are just as described—except that the violations are caused by these features as they occur on longer phrases, not in the lexical items. Thus

(18) *much honeys

is bad because honeys has been marked +count à la (iii) and (as the syntactic approach had already alleged in (8) above), much cannot be combined with +count on syntactic grounds. This can hold even though

(19) a. most honeys
b. most honey
c. much honey
d. most honey that comes from clover

are perfectly fine NPs. These examples show that in this theory, determiners like much and most can combine with a lexical noun (unspecified for +mass/+count) to generate an NP that does have one of the features, as in (19a),8 (19b), and (19c);

7 One aspect I will skip over here is whether +mass and +count exclude one another. The suggestion being made here assumes they do, but it could easily be modified to allow that some complex phrases could be syntactically marked with both +mass and +count.
8 Actually, in (19a), the unspecified-for-mass/count lexical noun honey is pluralized, yielding honeys—which is marked +count. As remarked in (8d), most can combine with singulars to produce +mass, but can also combine with plural +count (and the result will be +count).
and they can also combine with CNPs, where they will inherit the feature of the CNP, as in (19d). Of course, when the feature of the CNP violates the agreement feature of the determiner, as in (18), this generates syntactic ill-formedness. It is difficult to see how this could be accommodated in a theory of syntactic +mass/+count lexical features, except by having a stable of syntactic coercion rules that changed +mass to +count and conversely. But that’s just a bad syntactic idea. I think the examples of section 2.4, as well as the apparent viability of the present approach, show that the syntactic +mass/+count features just shouldn’t be associated with the lexical nouns, but rather ought to be ‘constructional’ in nature—introduced in the construction of larger phrases.

2.7.2 Lexical nouns are both +mass and +count

As discussed in section 2.3, semantic theories take the point of view that +mass and +count are ways of describing the semantic values of terms. If a term’s semantic value exhibits the properties identified in the first clauses of the criteria listed in (12) above, then we say that the term is +mass; if it exhibits the features identified in the second clauses of those criteria, then we say the term is +count. These semantic categorizations do not themselves feature into well-formedness considerations; rather, it is just a matter of brute semantic fact that certain semantic properties cannot be combined with other semantic properties. The attempt to do so results in a ‘semantic anomaly’.

The theory being developed here holds that the semantic value of every lexical noun contains all the values of which the noun is true. Thus, since a noun such as chocolate is true of some individual candies as well as of the stuff of which they are made, both of these meanings will be identified as parts of the semantic value of the lexical noun chocolate. For concreteness, I’ll employ the usual semantic value of terms identified as +count—namely, the set of things of which it is true (or a function on possible worlds that picks out the set of chocolate candies in that world). And for a mass meaning, I’ll adopt the view that it is a semi-lattice structure (atomic or non-atomic, depending on the noun) or an ‘ensemble’ or a mereology (atomic or non-atomic). There is no necessary commitment to these for semantic values in either the +count or the +mass case, although I do think they are plausible. So the semantic value of the lexical noun chocolate would be the union of the set of individual chocolate candies and the semi-lattice of chocolate.

Of course, we wish to preserve semantic compositionality. When the lexical noun chocolate is syntactically combined with the determiner a to form the NP a chocolate—and recall that this entails adding the syntactic feature +count to the NP—the semantic rule that corresponds to this syntactic combination has the effect of deleting the mass part of the meaning of the lexical chocolate. That is, the resulting semantic value of a chocolate now contains only the set of individual pieces of
chocolate candies, and not the semi-lattice of chocolate. In turn, this means that the
description of the semantic value of a chocolate obeys the semantic characteristics of
+ count.

This strategy is different than semantic ones prevalent in the literature. Standard
views take one or another of the + count or + mass meanings as basic, and then
expect the semantic rule to construct the other one upon demand. For example, if
the + mass notion of chocolate were taken as basic, then the semantic rule we have
been discussing would have to construct the correct meaning of ‘set of individual
pieces of chocolate candies’ out of the semi-lattice of chocolate. However, in the
strategy being advocated here, rather than trying to perform some sort of type-shift
or a coercion or a construction of a related meaning, all these values already are part
of the lexical meaning of chocolate. And the effect of the semantic rule—indeed, the
semantic rules for all the different syntactic combinations—is to delete some aspects
of the lexical item’s semantic value from consideration in the current syntactic
context. In this way, the meanings of lexical items are both + mass and + count.

I have already remarked that a great number (the vast majority?) of nouns have
both a ‘natural’ count meaning and a ‘natural’ mass meaning. These are the nouns
that I was calling ‘dual nouns’ (and their paired meanings ‘dualities’). Furthermore,
the universal grinders and packagers show that at least the non-abstract nouns
have these sorts of mass and count meanings already embedded within their
semantic values, needing only some appropriate context to become highly salient.
We discussed the example of beer, in whose extension we find not only the semi-
lattice of beer, but also individual servings of beer, standardized types of individual
servings of beer, kinds of beer, and perhaps other types of values as well. In the
present proposal, all these will be part of the semantic value of the lexical item beer.

In more general terminology, the proposal for lexical semantic value is this. Given a [− abstr] lexical noun N, its (extensional) semantic value, \( \mu(N) \), would be
(something like):

\[
(20) \quad \mu(N) = \{ N^o \cup N^m \cup N^s \cup N^ss \cup N^k \cup \ldots \}
\]

that is, the union of all the things of which it is true. (\( N^o \) represents the objects that
are N; \( N^m \) is the material that N is true of; \( N^s \) are the standard servings of N; \( N^ss \) are
the standard sizes of servings of N; \( N^k \) are the kinds of N; etc.)

The present proposal is a step of taking away some information from the lexicon:
all syntactic count/mass information appears only in more complex phrases, and
thus lexical nouns are neither + mass nor + count in the present proposal. But the
proposal is also a step towards adding more information in the lexicon: the semantic
values of all the different uses of a noun become part of the lexical semantic value.
The characterization of the semantic values of nouns can be said to be both + mass
and + count because typical nouns have both semantic notions true of (different)
parts of the semantic value.
Lexical nouns are both and neither +mass and +count

It may be the case that there are some lexical nouns that have no +mass meaning—*piece* or *hole* are perhaps examples. And there may be some that have no +count meaning—perhaps *stuff* is such a noun. The semantic value of these lexical nouns then would correctly be described as being just one of +count or +mass, and not both. But they would still be classified as neither +mass nor +count in the syntactic sense, because those syntactic features are introduced at the level of the CNP or NP.

Of course, it is not really true that lexical nouns are neither +mass nor +count in the same sense in which they are both +mass and +count. The former sense is syntactic, the latter sense is semantic. For clarity we should subscript or otherwise indicate which sense we are discussing, whenever the topic of whether some piece of language is or isn’t +mass/+count (we might use +mass$_{\text{syn}}$ vs. +count$_{\text{syn}}$ and +mass$_{\text{sem}}$ vs. +count$_{\text{sem}}$, for example).

2.8 Related proposals

The approach most closely related to the present one is in Allan (1980), especially his main motivation of ‘countability is a subcategory of the NP, not of nouns’. However, the general framework within which Allan is working is influenced heavily by generative semantics, and the syntactic features of complex phrases are sometimes generated from semantic features of its parts. For example, Allan thinks that there are ‘eight levels of countability’ that attach to bare nouns, and that these levels of countability ‘can be incorporated into well-formedness conditions on English grammar’. So this general picture does not have a separation of the syntactic and semantic notions of +mass/+count in the way that the present proposal does, and its concept of a semantics seems much more ‘impressionistic’ than is common in today’s semantic theories.

The theory of +mass/+count in Borer (2005) is a purely syntactic theory without any semantic aspect at all (although it does seem possible to graft a semantic theory onto the proposal). Like the present theory, Borer’s theory also has lexical items be marked as unspecified for +mass or +count, meaning of course that they do not have +mass$_{\text{syn}}$ or +count$_{\text{syn}}$. But unlike the syntactic aspect of the present theory—where these features are added to the syntactic description of a more complex phrase when it is constructed from the simpler parts—Borer’s theory has the addition of one or another of these syntactic features occur as part of the morphology of words, and not as part of the syntactic composition of syntactically larger phrases from smaller ones within the syntax of a sentence.

The theory offered in the seminal articles Chierchia (1998a,b) is at heart a semantic proposal. (‘The mass/count distinction concerns the extension of predicates’, Chierchia 1998b: p.355.) He offers a characterization of the languages of the world in terms of two ‘semantic parameters’: ±arg and ±pred, depending on whether the language allows its nouns (and NPs) to have the semantic value of an argument.
(for common nouns, this means to take the semantic value, Kind) or to be interpreted as a predicate (with the semantic value of a property). These interacting parameters classify the world’s language into three types; the possibility of \(-\text{ARG} / \text{PRED}\) is not seen as possible because this ‘would prevent NPs from having any interpretation at all’ (Chierchia 1998b: p. 353). In Chierchia’s hands, this outlook on the interpretation of nouns in semantically different types of languages has implications for the way the nouns have to be represented in the lexicon. In \(+\text{ARG} / \text{PRED}\) languages, ‘all nouns are going to be, in some sense, mass’. He also reasons that there will be no singular/plural distinction in such languages, and that numerals will be unable to modify nouns directly but will require a classifier. Mandarin is cited as such a language. Languages that are \(-\text{ARG} / +\text{PRED}\) ‘should disallow bare nominal arguments altogether. Moreover . . . some nouns will have a count extension, while others will have a mass one’, and French is cited as an example. In \(+\text{ARG} / +\text{PRED}\) languages, ‘lexical entries can either denote kinds or predicates . . . and their phrasal projections can be freely shifted back and forth’. (Much of the relevant difference between \(-\text{ARG} / +\text{PRED}\) and \(+\text{ARG} / +\text{PRED}\) languages concerns the type of bare nouns that are allowed, and the way they are to be interpreted.) +ARG lexical nouns will have mass denotations; +PRED lexical nouns will have a set of atoms as its extension, and hence be count.

The present proposal differs from this in three ways: first, it does not attempt to derive mass and count interpretations of nouns from more general semantic properties; second, it finds a place for \(+\text{COUNT}_{\text{syn}}\) and \(+\text{MASS}_{\text{syn}}\), and does not consider all \(+\text{MASS} / +\text{COUNT}\) interactions to be semantic; and third, it does not employ the tactic of assigning one or another of a \(+\text{MASS}_{\text{sem}}\) or \(+\text{COUNT}_{\text{sem}}\) to the lexical nouns and then deriving the other meaning by means of some type-shifting operator when the linguistic context calls for it.

Bale and Barner (2009a) propose that the semantic values of lexical items are ‘underspecified’ with respect to count/mass, by which they mean that lexical items do not have either the syntactic \(+\text{MASS} / +\text{COUNT}\) features nor can their semantic values be characterized by the semantic properties of \(+\text{MASS} / +\text{COUNT}\). But their rule that interprets such an underspecified lexical item as it occurs in ‘a mass noun syntactic frame’ is the identity function. So: in reality their lexical items are not underspecified: they have mass meanings, and all count meanings are constructed in the relevant syntactic frames. So, like the present proposal, Bale and Barner’s proposal removes the syntactic \(+\text{MASS}_{\text{syn}} / +\text{COUNT}_{\text{syn}}\) features from the lexicon, and introduces them ‘in construction’. But the semantic side of their proposal has all lexical meanings be \(+\text{MASS}_{\text{sem}}\), and the \(+\text{COUNT}_{\text{sem}}\) feature be coerced from them in the appropriate syntactic constructions. As I have remarked earlier, I think this type of coercion is quite difficult to describe with any accuracy, given the numerous different ways that \(+\text{MASS}\) and \(+\text{COUNT}\) meanings are related.

Chierchia (2010) is a development, or perhaps an alteration, of the views in Chierchia (1998a,b). In this work, the treatment of \(+\text{MASS} / +\text{COUNT}\) in the world’s
languages is again divided into three types. But now the divisions classify the
languages differently, and the emphasis on $\pm \text{arg} / \pm \text{pred}$ is dropped. (We look
at his classification a bit more in section 2.9 below.) The emphasis in his 2010 article
is on the vagueness of the interpretation of atomic parts of mass terms. Chierchia’s
general proposal is that all nouns have atomic parts, instances of the extension of the
noun such that its subparts are not in the noun’s extension. But for a number of
nouns, it is vague as to what the atomic parts actually are. In light of this vagueness,
Chierchia opts for the semantic technique of supervaluation to give the appropriate
interpretation of certain types of sentences involving these mass nouns. Although
this view concerning the vagueness of what are atomic parts of some mass nouns
could be added to the ‘mass part’ of the semantic value of some nouns in my
proposal, other aspects of Chierchia’s proposal—such as the classification of some
nouns as $+\text{mass}$ and others as $+\text{count}$, and the extensive use of type-shifting to
account for dual uses of almost every noun—are not a part of the present theory.

2.9 Cross-linguistic comments

Chierchia (2010) gives a very helpful three-way division of how various languages
deal with the $+\text{mass} / +\text{count}$ (syntactic) distinction. According to this division,
the world’s languages fall into one of the following three groups with regards to
$+\text{mass} / +\text{count}$. First, there are number marking languages, which have overt
number features that obligatorily appear on nouns. Here the $+\text{mass} / +\text{count}$
distinction applies to the nouns directly. (Most?) Indo-European languages, e.g.
English, are such languages. Next, there are classifier languages, which do not have
obligatory number marking on nouns (and arguably do not have a singular/plural
contrast at all on nouns). Lexical nouns in such languages have been viewed as
$+\text{mass}$ in some influential works;⁹ but there is a $+\text{mass} / +\text{count}$ distinction in
these languages that is active more generally, and for this reason it might be better to
view the lexical nouns as unspecified for $+\text{mass} / +\text{count}$). The classifiers in these
languages enforce the $+\text{mass} / +\text{count}$ distinction, but at the level of an entire
‘classified noun phrase’.¹⁰ (Most?) Asian languages (whether related or not), e.g.
Mandarin, Japanese, and Korean, are such languages. Finally, there are languages
that lack both obligatory number marking and obligatory classifier systems. Various
Amerindian languages, e.g. the Canadian Dene Sųłiné, various South American
languages, e.g. the Brazilian Karitianan and Yudja, and perhaps some Austronesian
languages are such languages. Some of these languages can plausibly be seen as
having a $+\text{mass} / +\text{count}$ distinction, albeit on somewhat different bases than the
foregoing languages. For others, it seems that there simply is no such distinction
reflected in the language—whether the distinction is seen as semantic or syntactic.

⁹ E.g. Krifka (1995), Chierchia (1998a,b), and others.
¹⁰ See Cheng and Sybesma (1999) for an influential article arguing this viewpoint.
The present proposal is aimed at number-marking languages. But it might be noted that the proposal has the effect of bringing number-marking languages more into line with classifier languages, in that the +mass/+count distinction happens at a level of syntactic complexity that is larger than lexical nouns: e.g. at the level of CNP and NP, in the present proposal. Thus, under the current proposal, lexical nouns are unspecified for +mass<sub>syn</sub> and +count<sub>syn</sub>, just as in the classifier languages, and instead these features are introduced ‘in construction’. The difference between number-marking languages and classifier languages thus lies simply in the differing syntactic constructions that introduce the syntactic features. Number-marking languages introduce the feature +count<sub>syn</sub> in pluralization and in combination with certain determiners, and they introduce +mass<sub>syn</sub> in combination with other determiners. Classifier languages, on the other hand, introduce +mass<sub>syn</sub> or +count<sub>syn</sub> in construction with the different classifier phrases. But the result is the same: it is the syntactically more-complex constructs that are +mass or +count, and not the lexical nouns.

This manner of looking at the ways that number-marking and classifier languages treat the +mass<sub>syn</sub>/+count<sub>syn</sub> features brings the two categories of languages very much closer together than Chierchia’s initial description would suggest. And indeed, much closer than most descriptions of the +mass/+count phenomena of the two language families has ever suggested. The languages that are neither number-marking nor classifier form a diverse group. In many of these languages it seems that there is no syntactic clue as to whether a sentence involving a noun should be understood as +mass or +count. And either semantic understanding is plausible when looking at just the sentence out of context. It is then up to ‘the context’ to determine which it is. But in such a case, it is not right to say that the language is employing such a distinction. Whatever the distinction among such language users, it is to be on a different basis than the way it works in number-marking and classifier languages (see Lima 2010a, Müller et al. 2006, Wilhelm 2008).

2.10 Further advantages

Section 2.8 outlined various proposals in the literature that have some similarities to the present scheme. But although there are these similar proposals, I think there are some advantages of the present idea over them.

For one thing, my view acknowledges the extensiveness of ‘dual meanings’: although there may be a few lexical items that can’t be used ‘massily’ — piece or hole (?) — and a few that can’t be used ‘countily’ — stuff (?) — almost all can be used in both ways. Not only can they be in fact used in both ways, in a very, very large number of cases both meanings are equally ‘natural’. These dual-use nouns form a primary reason to adopt the current proposal. Syntactically speaking, the only two alternatives to saying that such nouns are neither +mass nor +count are either to
proliferate entries in the lexicon (noun \( X \) with the syntactic feature \( +\) count and a duplicate entry for \( X \) with the syntactic feature \( +\) mass) or else to engage in syntactic type-shifting (where some syntactic rule changes the \( +\) count \( \text{syn} \) feature to \( +\) mass \( \text{syn} \), and conversely). Neither of these alternatives seems very appealing, especially when one emphasizes the naturalness of both meanings of the dual-use nouns. In the theory as presented, the \( +\) mass \( \text{syn} \) and \( +\) count \( \text{syn} \) features that are introduced in construction actually do some syntactic work: certain constructions are deemed syntactically ill-formed because of violations of the restrictions that these features enforce. This is in contrast to the theories of syntactic features that can always be converted from \( +\) mass to \( +\) count, and conversely. It is also in contrast to those semantic theories that locate the ‘oddness’ of certain constructions as a kind of semantic clash between \( +\) mass \( \text{sem} \) and \( +\) count \( \text{sem} \).

And against theories that have a unified lexical entry but mark ‘different senses’ of the noun as \( +\) mass vs. \( +\) count, there is no need in the present approach for any fancy semantic operations such as ‘grinding’, ‘packaging’, semantic type-shifting, or other ‘coercions’. Nor need there be any ‘magic’ in choosing or describing which of the different senses will be employed in some particular sentence. (Note that, in theories employing ‘different senses of the same word’, if one is describing a sentence that employs a dual-use noun, then to accurately state which of the different meanings is being utilized one needs to look to larger parts of the sentence that contain the noun. That is, it becomes necessary to look beyond the word itself and to the larger context in which it occurs. But that is a violation of semantic compositionality, and should be avoided if there is an alternative available.)

As against the approach of Allan (1980), there is no attempt to make ‘levels of countability’ be a syntactic phenomenon (which even Allan admitted were individual matters of taste). And as against Chierchia (1998a,b), the present theory does not claim that basic lexical meanings are \( +\) mass, not even in the classifier languages. And so there is no need to have ‘coercions’ that will generate \( +\) count meanings from them. And more globally, the theory I am offering does not make any use of Chierchia’s \( \pm \text{arg} / \pm \text{pred} \) background theory. As against Chierchia (2010), there is no necessity of saying that the notion of object vs. mass (what are the atomic parts of mass terms) is ‘vague’—although one could; and so there is no need for unusual semantic techniques like supervaluations.

As opposed to Bale and Barner (2009a), lexical values really are underspecified (and also, of course, overspecified) in my theory: they are not ‘really’ mass-meanings that are to be ‘coerced’ into count-meanings in certain environments. In general, and as opposed to Chierchia and to Bale and Barner, my approach is not a semantic approach—violating the \( +\) count or \( +\) mass agreement features is a syntactic error. In general, and as opposed to Allan and to Borer, there is (or can be) a legitimate semantics for the lexicon which works compositionally with the syntax, and one need not buy into any of the other parts of their theories.
Finally, it should also be emphasized that, on the semantic side, the present theory’s view that lexical nouns are both +mass and +count allows it to avoid the semantic coercion operations that convert one to the other. And there is no need to try to determine which of the meanings is ‘really’ basic and which one is ‘really’ a secondary coercion. Rather, associated with certain of the syntactic rules is a semantic rule that has the effect of deleting some of the semantic values of the lexical items. So, instead of trying to generate or create new senses out of old ones using semantic coercion rules, we have all the possible meanings stored in the lexicon and simply carve out the ones that are relevant in some syntactic case or other. This seems a much cleaner way to accommodate the set of semantic facts that surround the facts of the near universality of the +mass/+count alternations.

2.11 A final philosophical remark

In philosophy, much has been made about the difference between stuff and things, substance and substances, gunk and atoms. The relationship between these ontological distinctions and the linguistic distinction of +mass and +count was initially brought to philosophers’ attention in Strawson (1959), using the terminology of ‘sortal predicates’ and ‘feature placing terms’ for (some) +count and (some) +mass nouns, respectively. (Sortal terms are ‘terms that provide a principle for distinguishing and counting individual particulars’ while feature-placing terms ‘provide a fundamental basis that is presupposed even by sortal predications.’ Strawson’s examples of feature placing predications are There is water here; Snow is falling and the like, which makes it clear that he has some subset of mass terms in mind for this.)

I’m not opposed to such metaphysical distinctions, and indeed think they mark an important difference. But I am a skeptic about whether there is any useful information to be gained about the distinction by looking at the +mass/+count distinction in language—any language, even English, which apparently is Strawson’s idea of a logically and ontologically perfect language.11

At the very least, philosophical writers who wish to employ this sort of feature to bolster the philosophical conclusions ought to take a much wider group of languages into account than just English and its ‘number marking’ relatives.

---

11 Mei (1961) challenged Strawson’s account of ‘the description of “our” conceptual scheme’ by pointing to Chinese as a case where the sortal/feature-placing distinction doesn’t work as Strawson claimed. ‘Strawson exploits facts peculiar to languages like English… In Chinese, Strawson’s criteria are inapplicable’, Mei said. ‘Strawson’s silence [about other languages] must either mean he thinks that all other languages conform to English in these ways or that English is the paradigm of all languages.…. What justification can Strawson offer for this act of linguistic imperialism?’ Furthermore, ‘[Strawson employs] Aristotelian arguments based upon the peculiarities of English and its relatives…. [Strawson’s notion of] “assertive ties”, and subject-predicate “congruence” only works for languages with sufficient inflection.’
3

Aspects of individuation*

ELIZABETH COWPER AND DANIEL CURRIE HALL

3.1 Introduction

The relation between classifiers and plural marking appears, in some respects, to be one of complementarity in both distribution and function. Languages that make extensive use of classifiers typically lack obligatory inflection for number, while languages with systematic number marking often lack classifiers. Both types of morphology are implicated in counting, as illustrated in (1):

(1) a. san *(ben) shu [Mandarin]
   three cl book
   'three books'
   (Cheng & Sybesma, 1999: 514)

b. three book*(-s)

In the Mandarin example in (1a), the presence of the numeral san requires the presence of a classifier, but there is no plural marking. In the English counterpart in (1b), plural marking is obligatory, and there is no classifier. This pattern is not universal: languages with classifiers may also have some form of plural marking, as we will see later in this chapter, and some languages with plural marking (such as Hungarian) do not use it in combination with numerals. The purpose of this chapter is to shed light on both the pattern and some of the exceptions (real and apparent) by proposing a formal representation of plurality and classification as two separate aspects of individuation, the semantic property that characterizes count nouns cross-linguistically. Drawing on data from English, Mandarin, Cantonese, Armenian, Korean, and Persian, we will show how the differences among these languages can be reduced to a small number of differences in which features the language makes use of, which of those features can project as syntactic heads, and the status of non-projecting features as modifiers or head features.

* We would like to thank Nasrin Matini, Safieh Moghaddam, Lisa Cheng, Rint Sybesma, Mike Barrie, Gabriela Alboiu, and participants at the 2008 Bilingual Workshop on Theoretical Linguistics, at the Mass–Count Workshop, and at the 2009 meeting of the Canadian Linguistic Association for their helpful comments, questions, and judgements.
Under our analysis, it is not necessary that a language be characterizable, as a whole, as a classifier language or as a plural-marking language. Rather, classifiers and plural marking may coexist in a language as long as only one appears in any given nominal.

3.1.1 Theoretical Framework

We adopt a version of Distributed Morphology (Halle & Marantz 1993) in which features are privative, and are organized into dependency trees as proposed by Harley (1994). In any system of privative features, the absence of a property is represented formally simply by the absence of the corresponding feature, rather than by the presence of a negative value. However, as has been pointed out for phonological features by Hall (2007) and Dresher (2009), the relative scope of contrasts is crucial to determining whether the absence of a feature is meaningful. If a feature is absent from a context in which it could not be present, then its absence is not contrastive. For example, in many languages the feature Voice is not specified on sonorants—not because sonorants are voiceless, but rather because they are all voiced; because there is no voicing contrast in this domain, no feature marking is necessary, and the absence of the feature does not signify the absence of the corresponding property.

We also assume, following Wiltschko (2009), that morphosyntactic features can appear either as heads or as modifiers, and that this difference also has an effect on whether the absence of a feature is contrastive. If a feature is the head of a syntactic projection, or a morphosyntactic dependent of a head, then its absence is significant. For example, in the system of features proposed for English inflected by Cowper and Hall (1999), past tense is marked by a feature precedence dependent on the inflected head; an IP that lacks precedence is interpreted as contrastively non-past. A modifier feature, on the other hand, is like an adjunct, and its absence is non-contrastive. In Halkomelem, for example, Wiltschko (2009) argues that gender is a modifier feature on determiners: as shown in (2), the feminine determiner the can only be used with feminine nouns, but the unmarked determiner te is simply vague as to gender rather than contrastively non-feminine.

(2) Gender as a modifier feature on Halkomelem determiners (Wiltschko 2009: 40)

<table>
<thead>
<tr>
<th>a.</th>
<th>*Ílhtel the swíyeqe.</th>
<th>c.</th>
<th>Ílhtel te swíyeqe.</th>
</tr>
</thead>
<tbody>
<tr>
<td>eat</td>
<td>DET.FEM</td>
<td>man</td>
<td>eat</td>
</tr>
<tr>
<td>Intended: ‘The man is eating.’</td>
<td>‘The man is eating.’</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>b.</th>
<th>Ílhtel the slhálì.</th>
<th>d.</th>
<th>Ílhtel te slhálì.</th>
</tr>
</thead>
<tbody>
<tr>
<td>eat</td>
<td>DET.FEM</td>
<td>woman</td>
<td>eat</td>
</tr>
<tr>
<td>‘The woman is eating.’</td>
<td>‘The woman is eating.’</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

While Universal Grammar makes available a finite set of morphosyntactic features, there are a few ways that languages can differ parametrically in the use of these features. A feature may be used in one language, and entirely absent from another. If used in a particular language, a feature can appear as either a head feature or a modifier feature. Furthermore, the semantic dependencies encoded in feature geometries can be mapped to syntactic structure in different ways. If a feature $F$ is
semantically dependent on another feature $G$, then $F$ can appear as morphosyntactic dependent of $G$ on a single head, or $F$ can be the head of a separate syntactic projection that selects a complement marked with $G$.

As a starting point, we assume that Universal Grammar makes available at least the following features relevant to the semantic field of individuation and number (Cowper 2005b, Cowper & Hall 2009b):

(3) $\#$: individuated
   $>1$: plural
   $\text{CL}$: classified

The feature $\#$ encodes individuation, and a nominal from which $\#$ is contrastively absent is interpreted as mass. Plural ($>1$) is semantically dependent on $\#$, being interpretable only on count nominals. A count nominal from which $>1$ is contrastively absent is interpreted as singular. What we will represent here as the feature $\text{CL}$ is the morphosyntactic encoding of classification—a feature or class of features that specifies a unit of individuation. Like $>1$, $\text{CL}$ is semantically dependent on $\#$, but it elaborates individuation along a different dimension. $\#$ itself quantizes a nominal; $>1$ further specifies the number of quanta; $\text{CL}$ further specifies the type of quanta.

As stated above, the status of $\#$ as contrastive or non-contrastive depends on where in the syntactic structure the feature appears. If $\#$ heads a syntactic projection, then its absence is contrastive. If it appears as a modifier of another projection, then its absence is non-contrastive. As we will argue here, number features can occur on lexical noun heads, as modifiers on the Determiner head, or in the head of a separate inflectional projection, $\#P$ (originally proposed by Ritter 1992). The cross-linguistic variation in the position of $\#$ is reminiscent of work on gender features by Ritter (1993), who argues that gender is on the lexical noun head in Hebrew, but on the inflectional number head in Romance. Clarke (in progress) makes a similar argument for aspectual features: the event–state distinction is made in the verb phrase in Inuktut and Japanese, but in INFL in English (Cowper 2005a) and Greek (Kyriakaki 2006).

3.1.2 The mass–count distinction

It has commonly been assumed (see, e.g. Chierchia 1998b, 2009; Gillon 2009) that English makes a lexical distinction between mass nouns, such as *rice, water, mud*, and *furniture*, and count nouns like *cat, book, pea*, and *chair*. However, the classification is far from rigid. English count nouns can generally be used with a mass interpretation, and, conversely, English mass nouns can generally be used with a count interpretation. This has been accounted for by invoking semantic operations such as the Universal Grinder (Pelletier 1975), which converts count nouns to mass nouns, and the Universal Sorter (Bunt 1985a), which converts mass nouns to count nouns. Interestingly, there are some nouns, such as *furniture*, that behave syntactically like mass nouns, but which

---

1 We set aside the fact that many languages distinguish duals from singular and plural, and that there are languages that also distinguish trials or paucals. See Cowper (2005b) for some discussion of such systems.
cannot be converted to count nouns by the Universal Sorter. We will argue that in fact, the vast majority of English nouns are unspecified for the mass–count distinction, and that interpreting ‘mass’ nouns as count, or ‘count’ nouns as mass, involves not coercion from one category to the other, but rather the morphosyntactic filling in of features that are not specified on the nouns themselves. Partly in line with Borer (2005), we propose that these features originate in syntactic projections above N in the nominal phrase.

In contrast to English, Chinese has been argued (Chierchia 1998b) to have only mass nouns lexically. Since a numeral can appear in Chinese nominals only with a classifier, it has been proposed that it is the classifier that individuates the underlingly mass noun so as to make it countable. However, Cheng and Sybesma (1999) have shown that Chinese does indeed have a lexical mass–count contrast. Count nouns appear with classifiers, whereas mass nouns can be counted only if they appear with a different sort of element that they term a massifier. We follow this approach for Chinese, and extend it to English to account for exceptional nouns like furniture, which, we will argue, has a representation essentially the same as that of a Chinese count noun.

In the sections that follow, we will first show how our system of representations applies to English and Chinese, and then turn our attention to the question of complementarity between classification and number marking. Looking at cases in Armenian, Cantonese, Mandarin, Korean, and Persian, we will show that while >1 and CL may both be used in a single language, these two elaborations of # do not co-occur on a single nominal projection.

### 3.2 Taxonomy of English nouns

English has often been described as permitting ‘coercion’ between count and mass nominals. (See, e.g. Zwicky 2006, 2008 for an informal discussion.) Under this view, nouns may be underlingly count or mass, but the Universal Grinder (Pelletier 1975) can turn any count noun to mass, producing substance readings, and the Universal Sorter (Bunt 1985a) can make any mass noun countable, yielding interpretations that involve varieties, servings, or other quanta. Contrastive specification of the sort described in section 3.1.1 permits a simpler account of these phenomena.

#### 3.2.1 The regular pattern: Eels and tea

The behaviour of ordinary English nouns is illustrated in (4) and (5). Both a canonically mass noun such as tea and a canonically count noun such as eel receive count readings when they combine with numerals, the plural suffix, or the indefinite determiner a(n), as in (4); they receive mass readings in contexts such as those in (5).

(4)  
a. \([dp \; \{ #p \; te\]}\) of Sri Lanka are particularly nice.  
b. I’d like \([dp \; a \; \{ #p \; tea \} \}], please.  
c. I’d like \([dp \; \{ #p \; one \; tea \} \}], please.  
d. I’d like \([dp \; these \; \{ #p \; two \; te\]}\], please.  
e. My hovercraft is full of \([dp \; \{ #p \; e\}]]\).
f. There is \([\text{dp} \ \text{an} \ [\pi \text{ eel}]]\) in my hovercraft.
g. There is \([\text{dp} \ [\pi \text{ one eel}]]\) in my hovercraft.
h. \([\text{dp} \ \text{These} \ [\pi \text{ two eels}]]\) won’t leave my hovercraft.

(5) a. The cup was full of \([\text{dp} \ [\text{np} \text{ tea}]]\).
b. There is \([\text{dp} \ [\text{np} \text{ eel}]]\) all over my hovercraft.

We propose that the nouns themselves are not inherently specified as grammatically mass or count, although their lexical semantics may make one reading seem intuitively more likely than the other. Rather, the interpretation of a noun as count or mass depends entirely on the presence or absence of the grammatical # head. The nominals in (4) have structures like those in (6):

(6) a. \[
\begin{array}{c}
\text{DP} \\
\text{D} \\
\ #P \\
\ # \\
\ N \\
\end{array}
\]

\[
\begin{array}{c}
\text{a(n)} \\
\text{\{ tea } \\
\text{ eel } \\
\end{array}
\]

b. \[
\begin{array}{c}
\text{DP} \\
\text{D} \\
\ (the) \\
\ #P \\
\ # > 1 \\
\ N \\
\end{array}
\]

\[
\begin{array}{c}
\text{\{ teas } \\
\text{ eels } \\
\end{array}
\]

c. \[
\begin{array}{c}
\text{DP} \\
\text{D} \\
\ these \\
\ #P \\
\ two \\
\ # > 1 \\
\ N \\
\end{array}
\]

\[
\begin{array}{c}
\text{\{ teas } \\
\text{ eels } \\
\end{array}
\]
In (6a), the vocabulary item \textit{a(n)} spells out the combination of unmarked (i.e. non-plural) \# and unmarked (i.e. indefinite) D.\textsuperscript{2} In (6b) and (6c), a \# head specified with >1 is realized by the plural suffix -\textit{s}. We assume that numerals occupy the specifier position in \#P, as in (6c); the presence of a numeral thus requires individuation.

The nominals in (5) have the structure shown in (7):

\[
\text{(7)} \quad \begin{array}{c}
\text{DP} \\
\downarrow \\
\text{D} \\
\emptyset \\
\text{N} \\
\{ \text{tea} \} \\
\{ \text{eel} \}
\end{array}
\]

Here, the contrastive absence of the \# projection gives rise to a mass interpretation.

Under this view, there is no featural distinction between nouns like \textit{tea} and nouns like \textit{eel}. They are equally capable of appearing in count and mass nominals, and no coercion is required in either case.

3.2.2 Furniture, footwear, and equipment

While most English nouns follow this pattern, and can thus be treated as unspecified for individuation, there are some nouns that behave differently. \textit{Furniture}, for example, is sometimes described as a mass noun that cannot be coerced into a count reading—it can occur as a bare singular with an indefinite-quantity interpretation as in (8a), and it strenuously resists being combined with the indefinite determiner, the plural suffix, and numerals, as in (8b–8e). There are a few other English nouns, such as \textit{footwear} or \textit{equipment}, that share this pattern.

\[
\text{(8)} \quad \begin{array}{l}
\text{a. The room is full of furniture.} \\
\text{b. *I ordered a new furniture from Ikea. It has three knobs on the front.} \\
\text{c. *Of all the furnitures in the world, he had to pick Louis XV.} \\
\text{d. *I ordered three furniture(s) from Ikea.} \\
\text{e. *If there’s one furniture I can’t stand, it’s Louis XV.}
\end{array}
\]

In our system, it is not possible to say that \textit{furniture} is specified as mass in some way that cannot be overridden by coercion. For us, the mass interpretation arises from the contrastive absence of \#, and the absence of \# is not contrastive on English nouns, because \# in English is the head of a projection higher than N. Instead, we propose that the anomalous behaviour of \textit{furniture}-class nouns is due to the presence of \#; a nominal like (the) \textit{furniture} has the structure shown in (9).

\[
\text{(9)} \quad \begin{array}{c}
\text{DP} \\
\downarrow \\
\text{D} \\
\emptyset \\
\text{N} \\
\{ \text{(the) furniture} \}
\end{array}
\]

\textsuperscript{2} Here, we follow Cowper and Hall (2002), rather than assuming that \textit{a(n)} heads only a \#P.
Furniture spells out a structure containing both # and N, in which N projects. In this structure, # is a modifier (sensu Wiltschko 2009), rather than a separate syntactic head as in (6).

The presence of # on N reflects the semantic fact that furniture, footwear, and equipment are inherently individuated—they denote a set of individuated items, with the cancellable implicature that the set is non-singleton. Syntactically, the presence of the bare # feature both precludes the possibility of combining such a noun with a separate # projection, ruling out expressions like *furnitures, *a furniture, and *three furniture(s), and forces singular agreement (The furniture is numerous and varied rather than *The furniture are numerous and varied).

3.2.3 Cattle and livestock

The existence of nouns that spell out bare # suggests that we might also find nouns realizing # and its dependent feature >1. Such nouns should also not combine with a(n) or -s, but they should consistently trigger plural agreement. English does indeed have a small number of such nouns that conform to this expected pattern; one example is cattle, as illustrated in (10).

(10) a. Cattle are grazing in the meadow.
    b. *Cattle is grazing in the meadow.
    c. *A cattle is grazing in the meadow.
    d. *Cattles are grazing in the meadow.

A nominal such as the cattle has the structure in (11):

(11) D (the) NP # >1 N cattle

33 While the normal interpretation of (ia) below is that the room contains more than one article of furniture, the exchange in (ib) is perfectly well-formed:
As in (9), the presence of # on the noun in (11) encodes its inherent individuation, and makes it semantically ineligible to combine with a separate # head.

While the bundling of # with N as in furniture and cattle occurs in English only on a few exceptional lexical items, this pattern is more prevalent in other languages.

3.3 Chinese

3.3.1 Count nouns and classifiers

We assume that Chinese nouns are lexically categorized as mass or count, as argued by Cheng and Sybesma (1999). However, Chinese count nouns cannot actually be counted without a classifier, as shown in (12), repeated from (1a).

(12) a. san ben shu Mandarin
   three cl book
   ‘three books’ (Cheng and Sybesma 1999: 514)

   b. *san shu

We propose that count nouns in Chinese are structured like furniture-class nouns in English. They bear the feature # as a lexical property, and it is N, not #, that projects. (12b) is thus ungrammatical for the same reason that *one furniture is ungrammatical in English: there is no projection in the nominal phrase whose specifier can host the numeral.

According to Cheng and Sybesma (1999: 515), classifiers ‘name the unit in which the entity denoted by the noun naturally occurs’. We implement this with the structure in (13).

(13) CLP
    san ‘three’
    CL
    ben
    NP
    #
    N
    shu ‘book’

(i) a. We put the furniture in the living room.
   b. Q: Is there furniture in the living room?
      A: Yes, there’s a sofa.

See Doetjes (1996: 48–9) for a discussion of this type of noun, which she terms a ‘mass group.’

---

4 Doetjes (1996) makes essentially this proposal, though she does not provide a formal representation.
We take CL to be a feature that elaborates individuation. Essentially, Cheng and Sybesma’s insight is captured by saying that CL in Chinese is a syntactic head that projects independently, but is semantically dependent on #. The effect of merging the Classifier head with a count noun like shu is thus to create an individuated nominal projection analogous to the English #P. We assume that numerals appear in the specifier of CLP in Chinese, in parallel to their position in English in the specifier of #P. One salient difference between the Chinese CLP and the English #P is that CLP usually has a specifier, while #P frequently does not.\(^5\)

3.3.2 Mass nouns and massifiers

Chinese mass nouns cannot appear with ordinary classifiers. Instead, they appear with what Cheng and Sybesma (1999) call ‘mass classifiers’, or ‘massifiers’, as shown in (14). Count nouns can also appear with massifiers, as shown in (15).

(14) a. san ping jiu  [Mandarin]
    three cl-bottle liquor
    ‘three bottles of liquor’

b. san ba mi
    three cl-handful rice
    ‘three handfuls of rice’

c. san wan tang
    three cl-bowl soup
    ‘three bowls of soup’  (Cheng and Sybesma 1999)

(15) liang xiang shu
    two cl-box book
    ‘two boxes of books’  (Cheng and Sybesma 1999)

Unlike the classifiers discussed above, which merely name the unit of an already-individuated entity, massifiers both individuate the NP they apply to and name the unit of individuation. The unit frequently corresponds to a container or measure, as in the examples in (14) and (15). In addition, massifiers can be counted without adding another classifier. They thus resemble Chinese count nouns in that they contribute the feature # and describe an entity, and also resemble classifiers in that they can combine directly with numerals.

Cheng and Sybesma (1998) argue that massifiers are nouns that move from N to CL within CLP. For us, this gives the structure in (16).

\(^5\) In Cantonese, and to a slightly more limited extent in Mandarin, as discussed by Cheng and Sybesma (2005), a nominal can consist of only a classifier followed by a noun, as shown in (i).

(i) Go siupangjau fan gan gaau.  [Cantonese]
    cl child sleep asp nap
    ‘The child is having a nap.’  (Cheng 2008)
3.4 Plurality and classifiers

We have proposed that both >1 and CL are semantic dependents of #. However, it has been widely observed that languages that use classifiers seem not to use the singular-plural distinction, and vice versa. Ideally, this pattern should follow automatically from the representations of plurality and classification.

We propose that plurality and classification are different dimensions of individuation, just as location, time, and person are different dimensions of deixis. Ritter and Wiltschko (2005) argue that deictic anchoring of the clause is the core function of INFL, and that languages can differ as to whether that anchoring is temporal, spatial, or personal. They show that Blackfoot anchors clauses using personal deixis, and Halkomelem uses spatial deixis, while English uses the more well-known temporal deixis. Louie (2008) explores in detail the consequences of this difference for the representation of events and aspect in Blackfoot.

Returning to individuation, we can see the difference between English and Mandarin nominals as analogous to the difference between the INFLs of Blackfoot, English, and Halkomelem. English elaborates individuation using plurality, while Mandarin elaborates it using classification.

It is an empirical question whether the two ways of elaborating # are mutually exclusive, and whether any complementarity between them holds of an entire language, or only of a given nominal. In principle, it might be that languages can be classified as to whether they make use of classifiers or plurality, or it could be that nominals in a single language are divided into those with plurality and those with classifiers.

In the literature, languages with classifiers are often described as lacking ‘number’. From the data presented in the previous section, it seems that Mandarin does not make use of the grammatical feature >1 (although of course it has words that
lexically express particular numbers of entities), and English does not appear to make use of CL (although it has words that lexically express various kinds of units). Do any languages combine the two?

Armenian is reported to have both a classifier and an inflectional plural suffix, as illustrated in (17). The Armenian data below are from Borer (2005: 94–95); see also Bale and Khanjian (2009) for similar examples.

\[(17) \text{ a. Yergu had hovanoc uni-m. [Armenian]}
\text{two CL umbrella have-1SG}
\text{‘I have two umbrellas.’}
\text{b. Yergu hovanoc-ner uni-m.}
\text{two umbrella-PL have-1SG}
\text{‘I have two umbrellas.’}
\]

These two forms of individuation, however, apparently cannot be combined:

\[(18) \text{ *Yergu had hovanoc-ner uni-m. [Armenian]}
\text{two CL umbrella-PL have-1SG}
\text{‘I have two umbrellas.’}
\]

This suggests that while a single language may make use of both CL and \(>1\), a single nominal cannot. We now turn to a selection of cases where classifiers and plural marking have been proposed to co-occur within a single language and, more interestingly, within a single nominal. In each case we will see that a single instance of individuation (#) cannot be elaborated simultaneously by dependent features of plurality and classification.

### 3.5 Plural marking in Chinese?

#### 3.5.1 Cantonese di

A potential objection to the hypothesis that CL and \(>1\) cannot co-occur is based on the Cantonese classifier di. Di is often described as marking plurality, and is incompatible with a singular reading in some contexts. For example, Cheng (2009b) notes that in (19), ‘it is necessarily more than one sweater’:

\[(19) \text{ Wufei di laangsaam [Cantonese]}
\text{Wufei CL sweater}
\text{‘Wufei’s sweaters’ (Cheng 2009b)}
\]

---

6 As Bale, Gagnon, and Khanjian (2010, 2011) argue, the semantic contribution of the Armenian plural suffix -ner is not identical to that of plural marking in English. In Bale et al.’s account, -ner-marked nouns denote only non-singleton entities, whereas English plurals encompass both singleton and non-singleton members of a Linkian semi-lattice (Link 1983). We believe that our analysis of the morphosyntax of number and classifiers is at least broadly compatible with this treatment of cross-linguistic variation in the semantics of plurality.
However, *di* displays other characteristics that are less consistent with the notion that it spells out >1. It can be used with mass nouns without imposing a count interpretation, as in (20), and it cannot combine with numerals greater than one, as shown in (21).

(20) jat di seoi
    one CL water
  ‘some water’

  (Cheng 2009b)

    I buy PFV one CL bowl
  ‘I have bought a number of bowls.’

  (Cheng 2009b)

b. *Ngo mai zo saam di wun.
    I buyPFV three CL bowl
  ≈ ‘I have bought three numbers of bowls.’

  (Au-Yeung 2007)

This contrasts sharply with the English plural, which, as illustrated in (22), entails individuation and is compatible with numerals greater than one.

(22) There were two coffees, a tea, and five waters left on the table.

Au-Yeung (2007: 4) notes that *di* encodes a ‘fuzzy’ expression of quantity and has ‘a non-collective property that prohibits its countability. This property unspecifies how the referents denoted by the noun phrase group together and the grouping does not provide a shape or unit for counting’. We conclude that *di* is a massifier, and that it does not encode plurality. Rather, it indicates a very non-specific unit of individuation, one that is compatible with mass and count nouns alike, but that is not concrete enough to permit enumeration.

The nominals in (21) thus have the structure in (23). There is no syntactic impediment to inserting a numeral other than one in the CLP specifier position; the problem is semantic incompatibility between the vagueness of the unit specified by *di* and numerals greater than one.

(23)

```
(23)  CLP
       jat/’saam
       ‘one’/’three’
         CL
           N
             tN
tN
         #N
          di
          ‘vague amount’
          NP
            wun
            ‘bowl’
```
3.5.2 Mandarin -men

Another potential example of >1 in a classifier language is the Mandarin morpheme -men, which looks very much like a plural suffix. It appears on pronouns, as in (24), and also on nouns, as in (25).

(24) Mandarin pronoun paradigm:

\[
\begin{align*}
wo & \quad \text{‘I’} \\
ni & \quad \text{‘you (sg.)’} \\
ta & \quad \text{‘he/she’}
\end{align*}
\begin{align*}
\text{women} & \quad \text{‘we’} \\
nimen & \quad \text{‘you (pl.)’} \\
tamen & \quad \text{‘they’}
\end{align*}
\]

(25) Wo qu zhao haizi-men. \quad [Mandarin]

I will go find child-men

‘I will go find the children.’ \quad (Li 1999)

However, the distribution and interpretation of -men suggest that it is not an ordinary plural marker. It appears only on nominals referring to animate, and usually human, beings, such as those in (26):

(26) a. laoshi-men \quad \text{‘teachers’} \quad [Mandarin]

b. xuesheng-men \quad \text{‘students’}

c. pengyou-men \quad \text{‘friends’}

d. xiongdi-men \quad \text{‘brothers’}

e. jiemei-men \quad \text{‘sisters’} \quad (Li & Thompson 1989)

While the absence of -men on pronouns is contrastive, its absence on nouns is not. Whereas the pronoun wo in (24) is singular, rather than neutral as to number, the noun haizi in (27) is simply unspecified for plurality.

(27) Wo qu zhao haizi. \quad [Mandarin]

I will go find child

‘I will go find {the/some} {child/children}.’ \quad (Li 1999)

Furthermore, -men is incompatible with numerals (and with classifiers), as illustrated by the ungrammaticality of (28).

(28) *san ge xuesheng-men \quad [Mandarin]

three cl. student-men

Intended: ‘three students’ \quad (Li 1999)

Finally, the semantic contribution of -men appears to include more than just plurality. As pointed out by Li (1999), a nominal with -men is always interpreted as definite. Thus, -men cannot be used, for example, in existential constructions like the ones in (29).

(29) a. You ren. \quad [Mandarin]

have person

‘There is/are some person(s).’

b. Mei you ren. \quad [Mandarin]

not have person

‘There is nobody.’
Li (1999) argues that -men surfaces in D, not in #. She proposes that nouns can be marked with -men only if they move to D, and that the movement of N to D is blocked by an intervening classifier. While we do not adopt Li’s analysis in its entirety, we concur with her proposal that -men is associated with D, and that the presence of a classifier blocks movement of N to D, as shown in (30).7

Using Li’s analysis as a starting point, we propose that in Mandarin, the features definite, person (π), CL, and N can be syntactic heads. These features are shown in (31), with their dependent features and the vocabulary items that realize them. Our DefP corresponds to Li’s DP.

Li proposes an analysis that unifies the underlying representation of plurality in Chinese and English, so that both languages have a Number projection, while only Chinese has a Classifier projection. For us, while both languages make use of the feature #, this feature does not have the same syntactic status in the two languages.
As shown in (31a), the suffix -men spells out a Def head with dependent features anim(ate) and >1. Any nominal marked with -men will therefore be interpreted as animate, definite, and plural. The vocabulary items ta, ni, and wo spell out features at different levels in the person dependency structure in (31b): wo realizes speaker, ni is used for non-speaker discourse participants, and ta is used for non-participants. Mandarin pronouns have the structure shown in (32), in which a Def head takes a πP complement. If Def is specified with both anim and >1, then -men is inserted; otherwise, the Def head is not overtly realized.

(32) Mandarin:

<table>
<thead>
<tr>
<th>Structure</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>DefP</td>
<td>ta</td>
</tr>
<tr>
<td>Def</td>
<td>ta-men</td>
</tr>
<tr>
<td>πP</td>
<td>wo</td>
</tr>
<tr>
<td></td>
<td>wo-men</td>
</tr>
</tbody>
</table>

The proposal that -men spells out both animate and >1 correctly predicts an interesting quirk in the distribution of third-person pronouns: while ta ‘he/she/it’ can be used to refer to both animate and inanimate objects, tamen ‘they’ is used only for animates (Ng 1997). For plural inanimates, ta, not tamen, is used. Citing Chao (1968), Ng (1997: 8) says that ‘this non-use of tamen for inanimate things is the reason why a Chinese student of English will say These pears have spoiled; better throw it away.’

When -men appears with a lexical noun, the structure is as shown in (33). The plural feature on Def selects a count noun, since >1 is semantically dependent on #.

(33) Mandarin:

<table>
<thead>
<tr>
<th>Structure</th>
</tr>
</thead>
<tbody>
<tr>
<td>DefP</td>
</tr>
<tr>
<td>(dem)</td>
</tr>
<tr>
<td>Def</td>
</tr>
<tr>
<td>ANIM &gt;1</td>
</tr>
<tr>
<td>NP</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Examples:</th>
</tr>
</thead>
<tbody>
<tr>
<td>haizi-men</td>
</tr>
<tr>
<td>na-xie xuesheng-men</td>
</tr>
</tbody>
</table>

For completeness, the structure we propose for classifier phrases is shown in (34).
Mandarin:
Structure:

Examples: san ge haizi ‘(the) three children’
san ba mi ‘(the) three handfuls of rice’

While a DefP with no number feature can co-occur with CLP in Chinese, as in (35), a DefP containing the feature >1 cannot, as illustrated in (28), repeated below in (36).

(35) [DefP na [CLP san ben [NP shu ]]]
    ‘those three books’

It thus appears that the Chinese data are consistent with the hypothesis that while a single language may indeed have both plural marking and numeral classifiers, a given nominal projection will have at most one of the two. In the next two sections, we consider some possible counterexamples in Korean and Persian.

3.6 Apparent plural marking in Korean

Kim (2005) claims that in Korean, plural marking can co-occur with a classifier in a single nominal, as shown in (37).

(37) sey myeng-uy haksayng-tul
    three CL-gen student-PL
    ‘three (specific) students’

However, there is considerable evidence that the Korean ‘plural marker’ -tul is not, in fact, a true plural marker.

First, -tul is not obligatory in nominals with plural reference. A bare nominal can be interpreted as either singular or plural, as in (38). The interpretation of haksayng-i ‘student’ as singular or plural depends entirely on the context.
Second, \(-tul\) can appear on non-nominal constituents, as shown in (39).

\[(39)\]

\[
\begin{align*}
a. \text{Haksayng-tul-i wundongcang-eysel yelsimhi taliko-iss-ta.} & \quad \text{[Korean]} \\
\text{student-tul-nom playground-loc hard run-prog-decl} & \\
\text{‘Students are running hard at the playground.’} & \quad \text{(Kim 2005)} \\
\end{align*}
\]

b. \text{Nohi-tul oso-tul tul-o-tul w-a sul han can-tul ha-ko-tul ka.} \\
\text{you-tul quickly-tul enter-tul and wine one glass-tul drink-and-tul go} \\
\text{‘You come in quickly and drink a glass of wine and go.’} & \quad \text{(Lee 2000)}

Of particular interest is the fact that \(-tul\) appears on the noun can ‘glass’, in the nominal \textit{sul han can-tul} ‘one glass of wine’, without causing a contradiction. If \(-tul\) were an ordinary plural marker, it should be incompatible with ‘one’.

Indeed, Park (2008) has argued that \(-tul\) is not a plural marker, but rather a distributivity marker, whose meaning is similar to floated \textit{all} in English. It can appear with distributed predicates, and also with those collective predicates that have distributive sub-entailments. When a collective predicate has a distributive sub-entailment, every participant has the property of taking part in the action denoted by the predicate (Dowty 1987).

In (40), the individual members of the group denoted by the subject cannot be interpreted as taking part in ‘being a very big group’; only the collective can. \textit{-Tul} is therefore disallowed, and (40b) is ungrammatical.

\[(40)\]

\[
\begin{align*}
a. \text{Wuli-nun acwu khun tanchey-ita.} & \quad \text{[Korean]} \\
\text{we-top very big group-cop.decl} & \\
\text{‘We are a very big group.’} & \\
\end{align*}
\]

b. \text{*Wuli-tul-un acwu khun tanchey-ita.} \\
\text{we-tul-top very big group-cop.decl} \\
\text{Intended: ‘We are (all) a very big group.’} & \quad \text{(Park 2008: 283)}

In contrast, the predicate \textit{twule} ‘surround’ does have a distributive sub-entailment, and \(-tul\) is possible, as shown in (41).

\[(41)\]

\[
\begin{align*}
\text{Haksayng-tul-i sensayngnim-ul twule-ssassta.} & \quad \text{[Korean]} \\
\text{student-tul-nom teacher-acc surround-past.decl} & \\
\text{‘The students (all) surrounded a teacher.’} & \quad \text{(Park 2008: 283)}
\end{align*}
\]
Some predicates, such as ‘carry a box’, may be interpreted either distributively, or collectively with a distributive sub-entailment. In Korean, -*tul* is compatible with both readings of such a predicate, and a sentence like (42) is therefore ambiguous.

(42) Haksayng-*tul*-i sangca-lul wunpanhayessta. [Korean]
    student-*tul*-nom box-acc carry-past.decl
    ‘The students (all) carried a box.’ or ‘Each student carried a box.’ (Park 2008: 286)

Still other predicates, such as ‘be too heavy to carry’, can be interpreted either distributively or collectively, but the collective interpretation lacks distributive sub-entailment. In Korean, such predicates are disambiguated by -*tul*, which is compatible only with the distributive reading, as shown in (43).

(43) I kapang-*tul*-un naluke-ey nemu mukepta. [Korean]
    this bag-*tul*-top carry-to too heavy.decl
    ‘These bags are (all) too heavy to carry.’ (Park 2008: 287)

Park (2008) also argues that -*tul* appears in D/Q, higher than the Number or Classifier projections, and that it behaves like a strong determiner. A predicate nominal such as the one in (44) cannot be marked with -*tul*, and a -*tul*-marked nominal must take scope over negation, as in (45), and over the question operator, as in (46).

(44) Wuli-nun haksayng(+-*tul*-)ita. [Korean]
    we-top student-*tul*-cpl.decl
    ‘We are students.’ (Park 2008: 293)

(45) a. I kos-ey sensayngnim-un an kyesita.
    this place-loc teacher-top neg exist.decl
    ‘There is/are no teacher(s) in this place.’ or ‘The teacher is not in this place.’
  
    b. I kos-ey sensayngnim-*tul*-un an kyesita.
    this place-loc teacher-*tul*-top neg exist.decl
    ‘The teachers are not in this place.’ not ‘There are no teachers in this place.’ (Park 2008: 293)

(46) a. I kos-ey sensayngnim-i kyesiyeo?
    this place-loc teacher-nom exist.q
    ‘Is/are there a teacher/teachers in this place?’ or ‘Is the teacher in this place?’
  
    b. I kos-ey sensayngnim-*tul*-i kyesiyeo?
    this place-loc teacher-*tul*-nom exist.q
    ‘Are the teachers in this place?’ not ‘Are there teachers in this place?’ (Park 2008: 293)

---

8 Kwon and Zribi-Hertz (2004) argue, in a comparison of plural marking in Korean and in French, that -*tul* (their *deul*) is not inflectional, but lexical. They do not take a position as to where it appears in the syntactic structure; the important point for us is that they clearly distinguish it from the inflectional number marking found in French.
We conclude, following Park (2008), that Korean -tul is not a plural marker elaborating individuation. It is a type of distributivity marker, and can appear on non-nominal elements. When it appears in a nominal, it is part of the D system, not part of the number/individuation system. Korean is thus consistent with the proposal that classification and plurality are mutually incompatible elaborations of individuation.

3.7 Individuation in Persian

Ghomeshi (2003) and Gebhardt (2009) claim that nominals in Persian can contain both a classifier and a plural marker, as in (47):

(47) se ta gorbe-ha  
three CL cat-PL  
‘the three cats’

However, the Persian morpheme ta, glossed in (47) as a classifier, is crucially different from classifiers of the sort we have seen in Chinese, and the suffix -ha, glossed as plural, is not consistently associated with plural count nominals.

3.7.1 Ta: Not quite a classifier

The principal difference between Persian ta and classifiers in languages like Chinese is that ta does not actually classify: it is compatible with all nouns, and does not appear to indicate any specific unit of individuation. With canonically mass nouns such as čai (‘tea’), as in (48b), it produces a count interpretation, just as expressions like a tea or three teas do in English. Ta can also appear without an overt noun in anaphoric contexts such as the response in (48c).

(48) a. do ta deræxt  
two ta tree  
‘two trees’

b. Se ta čai, lotfæn.  
three ta tea please  
‘Three teas, please.’

c. Q: Čænd ta ænar xærid-i?  
how-many ta pomegranate bought-2sg  
‘How many pomegranates did you buy?’

A: Do ta xærid-æm.  
two ta bought-1sg  
‘I bought two.’

9 Except where otherwise noted, all Persian data are from Gebhardt (2009).
Ta must be preceded by an expression of quantity—either a numeral, as in (49), or
the interrogative/existential quantifier čænd, as in (50).

(49) a. *(se) ta pesær
   three ta boy
   ‘three boys’

   b. hær *(se) ta pesær
   each three ta boy
   ‘all three boys’

(50) a. Čænd ta ænar xærid-i?
   how-many ta pomegranate bought-2sg
   ‘How many pomegranates did you buy?’

  b. Čænd ta muš tuye zirzæmin hæst.
     some ta mouse in basement is
     ‘Some mice are in the basement.’

We propose that ta spells out a # head that obligatorily takes either čænd or a
numeral as a specifier, as in the structures shown in (51). Under this analysis, ta
codes the grammatical property of individuation itself. Ta thus differs from
Chinese count classifiers both in that it does not specify a particular unit of
individuation and in that it does not require its complement to be already individu-
ated. Rather than exhibiting a mass–count contrast like the one in Chinese, Persian
nouns, like most English nouns, are unspecified as to individuation, and so ta, like
the English indefinite article and plural suffix, can ‘coerce’ a count reading of a
canonically mass noun.

(51) a. #P
   čænd ‘how many’
   # ta
   ‘pomegranate’
   NP

   b. DP
   D hær ‘each’
   se three ta
   ‘boy’
   NP

Because nominals lacking ta may also be interpreted as count, as illustrated in (52), we
infer that Persian also has a phonologically null vocabulary item that can spell out #.

(52) Mina sib xord. [Persian]
   Mina apple ate,3sg
   ‘Mina ate apple/an apple/some apples/apples.’

We follow Ghomeshi, and assume that čændta is composed of a quantifier plus ta.
Since *ta* requires a specifier, the head of a # projection that lacks a numeral (or *čænd*) will always be null. When a specifier is present, some speakers always use *ta*, which is expected under the principle that more highly specified vocabulary items block less highly specified ones. For other speakers, *ta* is optional, being preferred in the spoken register but not the written one.

(53) Persian:
\[
\begin{align*}
\text{se } &\%\text{(ta) ostad} \\
\text{three } &\text{ta professor} \\
\text{‘three professors’}
\end{align*}
\]

For speakers who omit *ta* in writing, we can say that *ta* is further specified with a spoken register feature. This specification makes it incompatible with written contexts, thus allowing the null vocabulary item to spell out # instead. (See Cowper and Hall (2003) for a more detailed discussion of the role of register features in vocabulary insertion.)

3.7.2 -Ha: Not quite a plural marker

If *ta* spells out # rather than CL, then that fact by itself suffices to eliminate expressions such as (47) as instances in which a classifier and a plural marker appear in a single nominal. But if -*ha* really is a plural marker, then its co-occurrence with *ta* would still be surprising. In English, the presence of the suffix -*s*, which spells out >1, precludes the insertion of *a(n)*, which spells out # and (indefinite) D:

(54) a. *a cats
   b. *three a cats
   c. *a three cats

Cowper and Hall (2002) propose that it is not possible to insert vocabulary items separately spelling out both a dependent feature and a superordinate feature in the same hierarchy. If *ta* realizes # and -*ha* realizes the dependent feature >1, we would wrongly expect that -*ha* should block *ta*.

However, there is reason to believe that what -*ha* spells out is not, in fact, >1. First, unlike the English plural suffix, and like Mandarin -men and Korean -*tul*, Persian -*ha* is not obligatory on plural nouns; nominals without -*ha* can be interpreted as plural in an appropriate context:

(55) a. Muš tuye zirzæmin hæst. [Persian]
   ‘There are mice/there’s a mouse in the basement.’

   b. Mina sib xord.
   Mina apple ate.3sg
   ‘Mina ate apple/an apple/some apples/apples.’
Canonically mass nouns such as ‘sugar’ and ‘rice’ may receive various individuated interpretations when they appear with -ha, as illustrated in (56).

(56)  a. qænd-ha
      sugar-ha
      ’kinds/packets of sugar’

   b. Berenj-a-ro invar-o unvar næ-paš.
      rice-ha-ra here-and there NEG-throw.IMP
      ‘Don’t throw the (grains of) rice all over the place.’

However, adding -ha to a mass noun does not always force an individuated interpretation, as can be seen in (57), from Ghaniabadi (2009).

(57)  a. Ab-a qæt’-e.
      water-ha cut-is.3SG
      ’The water is shut off.’

   b. Bærf-a ab=šod.
      snow-ha water=became.3SG
      ’The snow melted.’

   c. Ru mase-ha či neveš-ti?
      on sand-ha what wrote-2SG
      ’What did you write on the sand?’

   d. Pa-m ru yæx-a liz=xord o oftad-aem o
      foot-1SG on ice-ha slippery=ate.3SG and fell-1SG and
      dæst-aem dær=raeft.
      hand-1SG door=went.3SG
      ’My foot slipped on the ice and I fell down and my hand got dislocated.’

One of our consultants notes that when -ha is used with mass nouns in this way, the nominal is interpreted as involving a large amount of whatever the noun denotes. For example, (57d) would be infelicitous if there were only a small patch of ice the size of a notebook. In that case -ha would be replaced by -e, which is normally used, in casual speech, for singular definite nominals.

We therefore propose that what -ha spells out is not specifically plurality, which entails individuation, but rather a more general property which can be understood, at least approximately, as ‘agglomeration’ (or AGGLOM for short).11 When it appears in combination with #, AGGLOM produces an interpretation that is effectively equivalent to a plural. Unlike >1, however, AGGLOM is not dependent on #, and when it appears without #, the result is understood as referring to a relatively large mass.

11 We choose this term as a means of expressing the notion of an augmented or non-minimal assemblage while remaining deliberately vague as to whether it is a mass or a group; the Oxford English Dictionary’s definition of agglomeration includes both ‘a mass formed by mere mechanical union or approximation’ and ‘a clustering or cluster’.
Another respect in which -ha is different from a regular plural marker—and similar to Mandarin -men—is that it appears to encode definiteness. Thus (58) cannot be interpreted as existential; (59) has only an equative reading; and in (60) the suffix -ra, which is obligatory on definite direct objects, cannot be omitted.

(58) Muš-ha tuye zirzæmin hæst-ænd. [Persian]
    mouse-ha in basement be-3PL
    ‘The mice are in the basement.’ NOT ‘There are mice in the basement.’

(59) Anha danešju-ha-ænd. [Persian]
    they student-ha-be.3PL
    ‘They are the students.’ NOT ‘They are students.’ (Ghomeshi 2003)

(60) Ketab-ha-ye jaleb-*(o) xund-æm. [Persian]
    book-ha-ez interesting-ra read.PAST-1SG
    ‘I read the interesting books.’

However, as noted by Ghomeshi (2003), nominals containing -ha can be made indefinite by the addition of the indefinite enclitic -i, as in (61), and -ha can appear on the nominal part of a compound verb, as shown in (62).

(61) a. Bæčče-ha-ye bahuš-i unja bazi=mi-kærd-æn. [Persian]
    child-ha-ez clever-INDEF there play=CONT-do.PAST-3PL
    ‘Clever children are playing there.’

    book-ha-ez interesting-INDEF read.PAST-1SG
    ‘I read (some) interesting books.’

(62) a. Bæhs-a=kærd-im. [Persian]
    debate-ha=do.PAST-1PL
    ‘We have debated often.’

   b. Kar-ha=kærd-im.
    work-ha=do.PAST-1PL
    ‘We have done many things.’

   c. Dærd-ha=kešid-im.
    pain-ha=pull.PAST-1PL
    ‘We have suffered repeatedly.’

Interestingly, our consultant reports that (62a) can be interpreted either as ‘We debated often’, i.e. on several occasions, or as ‘We debated a lot’, meaning at great length on a single occasion. Similarly, (62c) can mean either ‘We have suffered repeatedly’, or ‘We suffered a lot’. This range of interpretations is in some way parallel to the availability of both mass and individuated interpretations of -ha-marked nominals shown in (57) above.
While a comprehensive treatment of -ha would take us beyond the topic of this chapter, we believe that -ha can be accurately characterized as spelling out only **agglom**, and that in any case it does not realize >1. With respect to its role in the definiteness system, we tentatively propose the following: (1) -Ha itself is specified only with **agglom**. (2) In contexts such as (58), (59), and (61), -ha is inserted to spell out an instance of **agglom** that occurs as a modifier feature (*sensu* Wiltschko 2009) on D. (3) In Persian, the default interpretation of D is definite. A nominal containing a D head can be made indefinite by the addition of -i on a higher projection, possibly QP. Indefinite nominals containing neither -ha nor -i are either #Ps or bare NPs. (4) In compound verbs such as those in (62), -ha spells out an instance of **agglom** that is not attached to a D.

For the purposes of the question under consideration in this chapter, the only crucial aspect of this analysis of -ha is that it does not correspond to the feature >1. If this is correct, then we can conclude that a nominal like (47) represents neither the combination of a plural marker with a classifier nor the separate spelling-out of a dependent feature and its dominating node. Rather, the structure of (47) is as in (63).

(63)

```
(63)
       DP
          #P
            se
              'three'
            #
              NP
                ta gorbe
                'cat'

(63)
```

3.7.3 *Other possible classifiers in Persian*

While *ta* does not name any specific unit of individuation, and is therefore not a classifier in the relevant sense, there is a set of items that are more plausibly seen as classifiers. These include *jeld* (‘volume’), *livan* (‘glass’), *ghašogh* (‘spoon(ful)’), *dæste* (‘bunch’), and *mošt* (‘fist, handful’):

(64) do (ta) jeld ketab
    [Persian]  two *ta* volume book
    ‘two books’

(65) do (ta) livan čai
    two *ta* glass tea
    ‘two glasses of tea’

(66) se (ta) ghašogh næmæk
    three *ta* spoon salt
    ‘three spoonfuls of salt’
(67) pænj (ta) dæste gol
five ta bunch flower
‘five bunches of flowers’

(68) a. ye mošt berenj
one fist rice
‘a/one handful of rice’
b. se (ta) mošt berenj
three ta fist rice
‘three handfuls of rice’

These items can be used either in place of ta, as in (69a), or in combination with ta and/or -ha, as in (69b):

(69) a. do jeld ketab
   two volume book
   ‘two books’
b. pænj ta jeld ketab-ha
   five ta volume book-ha
   ‘the five books’

When one of these items appears together with ta, ta must come first, as in (69b). The order cannot be reversed, nor can either element be replaced by the other, as shown in (70):

(70) a. *do jeld ta ketab
   [Persian]
   two volume ta book
b. *do ta ta ketab
   two ta ta book
c. *do jeld jeld ketab
   two volume volume book

Like ta, these elements cannot appear without a numeral:

(71) *Jeld ketab xærid-æm.
    [Persian]
    volume book bought-1sg

These elements seem to be nouns, and indeed several of them are used independently as nouns. In the constructions shown here, we propose that they are merged in N, taking an NP complement, but that they move to # if they can. If # is occupied by another overt element such as ta, then they remain in situ. They are similar in some respects to the Chinese massifiers discussed above in section 3.3.2—nominal elements that individuate mass nouns by specifying a unit of measure. Unlike massifiers, however, they may remain in N and thus co-occur with the # head ta.
There is a small set of exceptions to the pattern just illustrated. According to our consultant, words like *kilo* and other purely measuring terms, and the word *næfær*, used with human beings and animals, cannot co-occur with *ta*. We also have no examples in which they co-occur with -*ha*.

(72) a. se (*ta) kilo gušt
    three ta kilo meat
    ‘three kilos of meat’

b. se (*ta) næfær kargær
    three ta person worker
    ‘three workers’

It would thus seem that in Persian, we find no reason to abandon our generalization that plural-marking and classification cannot co-occur in a single nominal. However, much work remains to be done in order to determine exactly what inflectional features are used in the Persian nominal, and how they map to syntactic projections.

### 3.8 Conclusions

In this chapter, we have been concerned with two aspects of grammatical individuation: numeral classification and plurality. We have proposed that individuation itself is a morphosyntactic feature, #, and that classification and plurality are two ways of elaborating the basic property of individuation. We represent these properties of nominals as features that enter into consistent semantic dependency relations, but which vary in their syntactic configuration. These features, part of a larger system of formal representations of the properties of nominal phrases, offer a precise and explicit way of characterizing the semantic and morphosyntactic roles of a range of morphemes, in various languages, whose meanings appear in some instances to overlap one another. For example, Persian *ta* is sometimes labelled a numeral classifier, and yet it shares only some of the properties of classifiers in Chinese: like them, it appears between a numeral and a noun, but unlike them, it does not appear to classify the noun or otherwise identify the unit of counting. In our system, this is explained by saying that *ta* is a # head, while numeral classifiers, narrowly defined, spell out a dependent feature of #. The similarity between *ta* and true numeral classifiers follows from the fact that *ta* spells out a feature that is also entailed by such classifiers. Similarly, the Mandarin suffix -*men* looks like a plural suffix because one of the features it realizes is the plural feature >1, but it differs from plural marking in English because it spells out other features as well, and because it is attached to DefP rather than #P.

While there is no obvious inherent semantic incompatibility between the two elaborations of individuation discussed in this chapter, the languages we have
investigated here seem never to allow a true numeral classifier and a true plural marker to co-occur on a single nominal. This is, as we noted above in section 3.4, reminiscent of Ritter and Wiltschko’s (2005) observations about the temporal, spatial, and personal manifestations of INFL in different languages. Whether the segregation of numeral classifiers and plural marking is a cross-linguistic universal, and if so, why, are questions that remain for further investigation; our system of featural representations provides formal tools with which to sharpen these questions and formulate possible answers.
Collectives in the intersection of mass and count nouns: A cross-linguistic account*

HEIKE WIESE

4.1 Introduction

In a language like English, which has systematic plural marking on nouns (henceforth 'plural language'), we find two major classes in the nominal mass/count domain: count nouns like cow or table, and mass nouns like beef or wood. A common observation for this domain is that these noun classes are distinguished by both their morphosyntactic behaviour and the status of their referents: nouns like cow pluralize, whilst nouns like beef do not (in the relevant reading), and nouns like cow refer to discrete objects, whilst mass nouns like beef refer to homogeneous substances.

This points to a one-to-one correlation of morphosyntactic and conceptual distinctions in the nominal mass/count domain. However, there is a third class of nouns, consisting of such nouns as cattle or furniture, that challenges this direct correlation. These nouns refer to discrete objects, and can accordingly appear in cardinal counting constructions, that is, in constructions that indicate an exact number of referents, a numerical quantity, rather than measuring a property like weight, volume, length, etc.\(^1\) Despite this 'count' behaviour, though, they do not

---

* For input and discussion, I would like to thank the audience at the Toronto Mass/Count Workshop, 2009, and two anonymous reviewers. For discussions and helpful intuitions on the Kurdish data, I am grateful to Sarkaut Zandi, Diler Assad, and Adel Zhia; for structural pointers for the Turkish data, I am indebted to Yazgül Şimşek.

\(^1\) Counting constructions with collectives are not two-term constructions, but typically involve a nominal classifier ('forty head of cattle / pieces of furniture'; see section 4.2.2.4 below). This might prima facie look like a measure construction. Crucially, though, the construction refers to a numerical quantity, that is, to an exact number of discrete objects, and these objects are identified by the collective noun. In contrast to that, measure constructions refer to properties like weight etc., and involve measure nouns (e.g., metres, kg) that identify different measure units and can be combined with different nouns. This will be treated in more detail in section 4.2.1.
pluralize and can appear as bare NPs without number marking (‘She bought cattle/ furniture’).

Hence these nouns establish an intersection between the major nominal mass/count classes in plural languages like English: on the conceptual side, they are in one class with count nouns, but on the morphosyntactic side, they go together with mass nouns. Following Bunt (1985a), I will refer to these nouns as ‘collectives’ for short. In section 2 below, I give a more thorough explication of the conceptual and syntactic distinctions that define this noun class and the other two classes involved in the nominal mass/count domain. For the time being, let us identify collectives as those nouns that refer to objects, rather than substances, but do not pluralize and can appear in bare NPs.

Another name used for this class of nouns in the literature is ‘object-mass nouns’. I prefer ‘collectives’ since the term ‘mass’ strictly speaking suggests a certain type of reference: reference to a homogeneous ‘mass’ entity, that is, a substance, rather than to a ‘count’ entity, that is, a discrete countable object. This, hence, might be misleading given that the class of nouns we want to distinguish here crucially does not refer to substances, but to countable objects.

Note, though, that the term ‘collective’ is sometimes also used for nouns that are excluded here, namely group nouns like family, committee, or indeed group, which behave like normal count nouns morphosyntactically (‘two families / committees’, ‘There is *group / a group’), and whose only special feature is that their referents are particular kinds of objects, namely sets – which is not something that is relevant for our discussion here. As far as the correlation of conceptual and morphosyntactic characteristics is concerned, these are regular count nouns: they refer to discrete objects (albeit a special kind of objects, namely, sets), and they pluralize systematically.

If we tease apart conceptual and morphosyntactic distinctions for our class of ‘collectives’, we get a pattern that goes diagonal to the one-to-one correlation characteristic for the other two nominal classes in a plural language like English. However, while this constitutes an exception in plural languages, it is the norm in so-called ‘classifier languages’, e.g. in Mandarin or Japanese, and also in some Indo-European languages, e.g. Persian (Farsi), where object-denoting nouns are collectives as a rule. Figure 4.1 brings the two cross-linguistic options together (‘SYN’: syntactic system; ‘CS’: conceptual system). In this graphic, as in the following discussion, I call nouns such as cow or table, which systematically pluralize, ‘plural nouns’, and use the traditional term ‘mass nouns’ for nouns like beef or wood that do not pluralize, but can appear in bare NPs, and refer to homogeneous substances.

---


3 In addition, my usage of the term ‘collectives’ does not cover pluralia tantum such as scissors, trousers, or glasses (in the sense of ‘binoculars’), which appear only in their plural form, even though they might refer to single objects.
There are two implications that this pattern suggests. For one, the different options to combine mass/count features from the syntactic and the conceptual system emphasize the linguistic arbitrariness of this correlation: they indicate a degree of liberty in combining the two levels that cannot be captured by one-to-one mappings between syntax and the conceptual system, but rather requires a mediating level in the linguistic architecture. In what follows, I am going to identify this as the level of grammatical semantics. Secondly, the cross-linguistic availability of collectives suggests a general mass/count distinction that is independent of syntactic plural marking in a language—and thus further emphasizes the autonomy of syntactic and conceptual distinctions in this domain. Note that (a) the option on the right brings together such typologically different languages as Mandarin, Japanese, and Persian, and (b) collectives appear in both configurations, albeit with different prominence. This suggests that the boundaries here are not between languages, but rather between noun classes: it suggests that languages differ not so much as to whether they have collectives (and classifier constructions) at all, but rather as to whether collectives are the dominant class of object-denoting nouns or not.

In what follows, I am going to elaborate on this picture. I will first give a definition of the syntactic and conceptual distinctions involved, and discuss the status of number markers and numeral classifiers from a cross-linguistic point of view (section 4.2). I will then provide further evidence for the independence of syntactic and conceptual classifications drawing on intra- and interlinguistic variation in mass/count correspondences (section 4.3). To account for the different options of correlating the two levels within a linguistic architecture, I will propose semantic representations that mediate between syntactic and conceptual distinctions (section 4.4). The final section (section 4.5) will then provide the conclusions from our discussion.
In order to characterize collectives as a nominal class apart from plurals and mass nouns, we have to define mass/count distinctions at two levels: the syntactic and the conceptual level. In this section, I propose definitions for this; the next section then presents data from linguistic variation involved in dialectal and diachronic developments and in semantic coercions, supporting our distinction between these levels.

4.2 Conceptual and morphosyntactic distinctions in the mass/count domain

In order to characterize collectives as a nominal class apart from plurals and mass nouns, we have to define mass/count distinctions at two levels: the syntactic and the conceptual level. In this section, I propose definitions for this; the next section then presents data from linguistic variation involved in dialectal and diachronic developments and in semantic coercions, supporting our distinction between these levels.

4.2.1 Conceptual distinctions

On the conceptual level, the relevant distinction is that holding between substances and objects: with respect to this level, nouns differ as to whether their referents are conceptualized as homogeneous substances or as (sets of) distinct objects. Following Prasada's (1996) explication of this distinction, I propose the following definitions to account for this:

(1) **Substances** are conceptualized as homogeneous entities whose structure is considered arbitrary.

**Objects** are conceptualized as (consisting of) individual entities whose structure is considered non-arbitrary.

This captures the traditional criterion for substances, which states that they have no minimal parts, without running into a problem if, say, we consider the molecular level where we might want to maintain minimal parts after all, given that the definition targets conceptualization, not ontology: referents of substance nouns are not necessarily totally homogeneous, and so of arbitrary structure, but are treated as if they are (cf. also Bunt 1985a).

With respect to this distinction, collectives and plural nouns form one class: their referents are individual, countable objects whose structure is non-arbitrary. To use a well-known example from the philosophical literature, one can divide the referent of water at random and the parts one gets will still be water, while the referent of cows cannot be divided like this—and the same holds for cattle: here, the minimal parts are individual animals, not, say, half a cow and two hooves of another cow. Accordingly, while one cannot count substances, but only measure them, one can count the referents of cows as well as those of cattle: in both cases, one counts whole animals, but not, say, half a leg.

---


5 Cf. also Gillon (1999), who emphasizes that nouns like cattle do not provide a full divisibility of their referents, and Bunt (2006), who notes that collectives and plural nouns do not differ at the level of meaning. Rothstein characterizes both noun classes as 'naturally atomic' and emphasizes that collectives like furniture or footwear denote 'sets of inherently individuable entities' (2010: 354). A different view is held by Chierchia (1998b: 347), who assumes that collectives and mass nouns are lexically pluralized and claims that
Note that what is crucial here is the distinction between grammatical behaviour and reference. Words such as *cattle* or *furniture* cannot appear in counting constructions without a classifier such as *head* or *piece* (see 4.2.2.4 below). This does not mean that cattle or furniture (= the objects) cannot be counted. It is perfectly natural to say, e.g. ‘Please count the furniture in the sitting room’. ‘Furniture’ refers to objects that come with a natural unit, that is, we can always say what counts as one, and whether there is one or two. This is what makes it possible for the noun ‘furniture’ to appear in counting constructions (as opposed to measure constructions), where the classifier does not add any content: ‘two pieces of furniture’ crucially does not refer to *pieces*, say, the leg of a table, but to two minimal instances of the concept *furniture*, e.g. a table and a chair. Hence, while ‘furniture’ cannot be combined with a cardinal numeral directly (‘‘two furniture’’) and does not pluralize (‘‘furnitures’’),\(^6\) it does refer to countable entities. These nouns are not uncountable (contra Wiltschko, this volume), in two respects: the nouns themselves are countable, of course (one noun, two nouns), and, more to the point, their referents are countable as well (e.g. one can count the furniture in a room). In their grammatical behaviour, they are like mass nouns in some respects, but accounting for this with a term like ‘(non)countability’ would fail to distinguish conceptual/referential and grammatical phenomena.

Hence, the classification of nominals as denoting countable objects is independent of their use with number markers: it encompasses collectives and plural nouns alike. This is in accordance with findings reported in Li, Ogura et al. (2009) from a study with English- vs. Mandarin-speaking toddlers, who show that children’s conceptual distinction of sets of one vs. many does not depend on the availability of syntactic plural marking in a language.

The conceptual allegiance of collectives with plural nouns—and thus the independence of the conceptual distinction from the syntactic one—is further highlighted by studies conducted by Barner and Snedeker (2005, 2006), who elicited quantity judgements for collectives, plural nouns, and mass nouns. Subjects had to compare scenes and say ‘Who has more… e.g. silverware (collective), shoes (plural noun), or toothpaste (mass noun)?’ Each scene showed one large instance of the nominal referents, for instance a large piece of silverware (e.g. a large fork), a large shoe, or a large mound of toothpaste, respectively, and one group of several small instances, that is, several small forks, shoes, or mounds of toothpaste, respectively. While the small instances were higher in number, the larger instances made up the bigger overall amount. Both children and adult participants chose the higher number of individual objects for both plural nouns and collectives, but the larger overall

\(^6\) Cf., however, the data in (12) below, which provides evidence for nouns such as *furniture* to be pluralized etc. in colloquial English. This further underlines my point: the referential countability that I emphasize here, can lead to grammatical ‘count’ behaviour in such cases.
amount for mass nouns. Hence, in experiments targeting referents’ conceptualization as countable objects vs. homogeneous substances, collectives, again, patterned with plural nouns.

In the same vein, Inagaki and Barner (2009) show that in Japanese, that is, in the absence of a systematic class of plural nouns, speakers also base their quantity judgements on the overall amount for mass nouns, but on the numerical quantity (= the higher number of individual objects) for collectives. This means that in a language that does not have syntactic plural marking and where nouns are generally not marked for number, we nevertheless find two nominal classes, ‘mass nouns’ and ‘collectives’, that are distinguished by the status of their referents—a distinction that is also reflected, for instance, by their distribution in counting vs. measure constructions.7 The difference between these two kinds of constructions is based on an important ontological distinction, namely that of measurement and counting:8 counting assesses numerical quantity (‘how many?’), while measurement assesses other properties than numerical quantity, e.g. weight, length, volume, temperature, etc. Accordingly, counting does not involve additional props, but maps numbers directly onto elements of a set, the counted objects. In contrast to that, measurement crucially involves additional measure units (kg, metres, °C, etc.). Since they do not relate to numerical quantity, measurement constructions are open to object- and substance-denoting nouns alike (‘four kg of gold’ as well as ‘four kg of apples’ and ‘four tons of furniture’), while counting constructions allow only object-denoting nouns.

4.2.2 Morphosyntactic distinctions

4.2.2.1 The notion of transnumeralty

On the morphosyntactic level, allegiances are reversed, and collectives are in the same class as mass nouns. Using a concept going back to Greenberg (1974a), I account for the relevant morphosyntactic feature by a definition of ‘transnumeral’ nominals, and call their counterparts accordingly ‘numeral nominals’.9

(2) In transnumeral nominals, number marking is not compulsory to indicate reference to more than one instance of the nominal concept; the nominal can occur in argument position as a bare non-plural noun phrase.

---

7 Cf. Li et al. (2008) on a developmental study indicating the distinction of the two noun classes in Mandarin, and Cheng and Sybesma (1998), who emphasize the distinction of numeral classifiers and measure words that occur in the respective numeral constructions.

8 Cf. Wiese (2003) for a detailed discussion and definition of these two different kinds of number mappings.

9 In addition to ‘transnumeral’, one also finds some other terms in the literature. For instance, Corbett (2000), in his discussion of number marking in the world’s languages, mentions ‘transnumeral’, but chooses the term ‘general number’.
In *numeral nominals*, number marking is compulsory to indicate reference to more than one instance of the nominal concept; the nominal does not occur as a bare, non-plural noun phrase in argument position, but is combined with morphological number markers or an article.

Following this definition, collectives and mass nouns are transnumeral, while plural nouns are not. The term ‘transnumeral’ refers to the fact that these nouns transcend number distinctions: if using a collective, one does not commit oneself as to whether its referent is a singleton or a set of more objects, whereas with a numeral nominal, one always has to identify whether one is talking of one or more. E.g., a sentence with a collective like ‘Karen bought furniture’ can refer to a situation where she bought one piece of furniture (e.g. an armchair) just as well as to one where she bought several pieces (e.g. a sofa, four chairs, and a table), but when using a numeral noun, e.g. *armchair*, one has to make clear whether she bought one (*an armchair*) or more than one (*armchairs*).

### 4.2.2.2 Number marking for numeral nouns

The main syntactic difference between transnumeral and numeral nominals is, then, the obligatory occurrence of number marking in numeral nominals. This can be realized morphologically or lexically. (3) gives cross-linguistic examples for morphological singular and plural marking:

\[
\begin{align*}
\text{a. } \text{drzewa} & \rightarrow \text{drzewo} & \text{[Polish: a tree—trees]} \\
\text{tree-} & \text{tree-} \\
\text{sg} & \text{pl} \\
\text{b. } \text{kitabu} & \rightarrow \text{vitabu} & \text{[Swahili: a book—books]} \\
\text{sg-book} & \text{pl-book} \\
\text{c. } \text{kitāb} & \rightarrow \text{kutub} & \text{[Arabic: a book—books]} \\
\text{book}_{\text{sg}} & \text{book}_{\text{pl}} \\
\end{align*}
\]

If we look at the English translations, we find that the corresponding singular forms involve the indefinite article. In languages like English, which have developed a systematic indefinite article (usually diachronically based on the numeral ‘one’), this article can be regarded as a lexical singular marker: it allows an expansion to a full, referential NP and marks the reference to a single entity, while the bare noun without number markers is neither specified for singular nor for plural reference yet, but serves as a basis for lexical singular and morphological plural markers alike:

\[
\begin{align*}
\text{d. } \text{a book} & \rightarrow \text{books} & \text{[English: lexical singular and} \\
\text{sg book} & \text{book-} & \text{pl morphological plural marking]} \\
\end{align*}
\]

\[10\] Note that on this view, we have to assume an implicit singular marker in constructions with the definite article.
We can account for this by assuming that for numeral nouns, the NP provides the lexical element that contributes the predication for the semantic representation, while a functional layer ‘TermP’ above NP, that corresponds, on the semantic level, to an instantiation function from predicates to terms, organizes number marking and maps them onto full, referential phrases that can appear in argument position.\textsuperscript{11} The head of TermP, then, contains either lexical number markers, such as the English indefinite article, or else a functional element that corresponds to morphological number markers in the head of the complement NP, such as English plural suffixes or singular and plural affixes in languages like Swahili. For transnumeral nouns, the expansion from NP to TermP is achieved without explicit elements, since transnumeral nouns can appear in argument position in their bare form, without number markers. Under a TermP account, we can account for this with a phonologically empty functional element which serves as the head of transnumeral TermPs. Fig. 4.2 shows the syntactic layers, including QP and DP (as additional functional layers above TermP), and the corresponding semantic processes:

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{figure4.2.png}
\caption{A functional layer TermP above NP}
\end{figure}

Note, though, that transnumeral nouns can be combined with number markers as well. This is captured in the definition above by characterizing number marking for transnumeral nouns as ‘not compulsory’. This can mean that it is not available for these nouns at all, as in English, but it can also mean that number marking is optional for them. This is frequently the case in languages with predominantly transnumeral nouns, that is, in languages where object-denoting nouns are generally collectives, rather than plural nouns, e.g. Mandarin or Persian.

4.2.2.3 Transnumeral number marking Transnumeral number markers differ from those for numeral nouns both syntactically and semantically: they are syntactically optional and thus serve as modifiers rather than heads and, corresponding to this, on the semantic level they indicate emphasis on a small or large amount. This semantic

\textsuperscript{11} Wiese (1997). Cf. also Ritter (1992), who assumes a functional layer NumP above NP that accounts for nominal number marking.
contribution can be characterized as ‘restriction’ (singular) and ‘amplification’ (plural),\(^{12}\) rather than the cardinality quantification as ‘one’ (singular) or ‘more than one’ (plural) that numeral number markers provide. (4) through (6) give some examples for transnumeral plural marking on collectives:

\[
\begin{align*}
(4) & \quad \text{a. mehmān} & \quad \text{[Persian]}^{13} \\
& \quad \text{guest} & \quad ‘\text{a guest / guests}’ \\
& \quad \text{b. mehmān.hā} & \quad ‘\text{a lot of guests / all kinds of guests}’ \\
(5) & \quad \text{a. háizi} & \quad \text{[Mandarin]}^{14} \\
& \quad \text{child} & \quad ‘\text{a child / children}’ \\
& \quad \text{b. háizi.men} & \quad ‘\text{several children, not only one}’ \\
(6) & \quad \text{a. cigaretta} & \quad \text{[Hungarian]}^{15} \\
& \quad \text{cigarette} & \quad ‘\text{a cigarette / cigarettes}’ \\
& \quad \text{b. cigaretták} & \quad ‘\text{several cigarettes / scattered cigarettes}’
\end{align*}
\]

\[(7)\] illustrates transnumeral singular marking for collectives with a Persian example:\(^{16}\)

\[
\begin{align*}
(7) & \quad \text{a. ketāb mixānām} & \quad \text{[Persian]} \\
& \quad \text{book read.1SG} & \quad ‘\text{I am reading \{a book / books\} / I am doing some reading.}’ \\
& \quad b. ketāb.ī mixānām & \quad \text{[Persian]} \\
& \quad \text{book- SG read.1SG} & \quad ‘\text{I am reading one book / a particular book.}’
\end{align*}
\]

Transnumeral number marking seems not to be principally restricted to languages with overall transnumeral nouns, like Mandarin or Persian, but can in some cases also occur for collectives in plural languages, e.g. in German. (8) provides evidence for this with data from Bavarian,\(^{17}\) where nouns like *Gäid* ‘money’ and *Gmias* ‘vegetables’, which are collectives in Bavarian (as are their standard German counterparts, *Geld* and *Gemüse*), are combined with the lexical singular marker we find in German, namely the indefinite article, with a semantic effect comparable to that noted for Persian in (7b):

\[12\] This goes back to Hincha’s (1961) work on Persian number marking. For semantic representations of transnumeral vs. numeral number markers, cf. Wiese (1997).

\[13\] Hincha (1961).

\[14\] Data from Kaden (1964:106).

\[15\] Data from Mikesy (1978:59).

\[16\] Data from Hincha (1961).

\[17\] Data from Merkle (1986).
(8) a. à Gäid [German: Bavarian]
    a money – ‘some money / a little bit of money’

b. à Gmias
    a vegetable – ‘some / a few vegetables’

Since transnumeral number marking does not identify an exact number, but rather indicates a large or small amount,\(^\text{18}\) it is not restricted to collectives, but can also occur with mass nouns, that is, with nouns referring to homogeneous substances, rather than objects. (9) gives an example from Persian that illustrates such transnumeral singular and plural marking for a mass noun, while (10) provides examples for collectives and mass nouns with optional plural markers in Turkish:\(^\text{19}\)

(9) āb – ‘water’ [Persian]
    ābī (sg) – ‘some water / a certain amount of water’
    ābhā (pl) – ‘a lot of water’

(10) a. Bu işte büyük para(lar) var. [Turkish]\(^\text{20}\)
    this business.loc big money.pl exist
    ‘There’s big money involved in this business.’

b. Kan(lar)a bulanmıştı. [Turkish]\(^\text{21}\)
    blood.pl.dat cover.refl.pfv.pst
    ‘He was covered all over in blood.’

c. Su yine keşilmişti. [Turkish]\(^\text{22}\)
    water again cut.pass.pst
    ‘The water is cut off again.’

c.’ Sular yine keşilmiş. [Turkish]\(^\text{23}\)
    water.pl again cut.pass.pst
    ‘All the water has been cut off again.’

This further highlights the different status of transnumeral number marking when compared to number marking for numeral nouns in plural languages. A phenomenon similar to transnumeral singular marking might be the combination of Mandarin mass nouns with the numeral yi ‘one’ + classifier xie, as described by Cheng (this volume). Wiltschko (this volume) discusses similar cases of pluralization

\(^{18}\) As (7b) above shows, in combination with collectives, this restriction to a small amount can suggest a single instance if the context supports this (cf. Hincha (1961)’s seminal work on Persian number marking for a more detailed discussion).

\(^{19}\) For a detailed discussion of transnumeral nouns and number marking in Turkish cf. Schroeder (1999).


\(^{22}\) Schroeder (1999: 209).

\(^{23}\) Schroeder (1999: 209).
for mass nouns in Halkomelem and Blackfoot. Note, though, that the semantic contribution of these number markers might be different from the ones discussed here, since they can occur (with object-denoting nouns) in counting constructions. Interestingly, there does not seem to be a semantic difference between constructions with and without number markers; Wiltschko (2008) states (for Halkomelem counting constructions) that ‘To the best of my knowledge, there is no meaning difference associated with the presence of the plural marker’. This might indicate a phenomenon of language change, namely a stage where number markers move from transnumeral to numeral status.

There are also instances from other languages with predominantly transnumeral nouns, for number markers that become similar to numeral ones, or at least move in that direction: number marking for collectives can interact with animacy in a way that nouns with referents higher on the animacy scale are more often pluralized, such that plural marking becomes less marked and, semantically, adds less emphasis on a large amount. An example for this is modern Turkish. However, nouns are still dominantly transnumeral in this case: they can expand to full referential TermPs without number markers, and, most notably, they cannot be pluralized in cardinal counting constructions; cf. (11a) vs. (11b):25

(11)  

(a) Çocuklar oynuyorlar.  
   child.pl play.pl  
   ‘The children are playing.’

(b) Üç (*Çocuklar / Çocuk) oynuyor(lar).  
   three child.pl / child play.pl  
   ‘There are three children playing.’

This second feature reflects a central characteristic that distinguishes transnumeral number markers from numeral ones, and is preserved in transnumeral languages such as Turkish, where number marking has become more general for object-denoting nouns: as noted above, transnumeral number, in contrast to numeral number, does not provide a cardinal quantification (‘one’ vs. ‘more than one’), but indicates an unspecified large or small amount, and this would clash with the exact cardinality provided by the numeral in counting constructions. As a result of this, numeral nominals are pluralized in counting constructions, whereas collectives are not; they obligatorily appear in their bare form. As mentioned above, these counting constructions then usually involve an additional element adjacent to the cardinal, namely a numeral classifier.26

26 In some languages, higher cardinal numerals sometimes behave like nouns, in which case they might take singular genitive complements (cf. Menninger 1958, Hurford 1975, Wiese 2003). While this, of course,
4.2.2.4 **Numeral classifiers in counting constructions with collectives** The occurrence of numeral classifiers further underlines the distinction of numeral and transnumeral nouns and number marking. Numeral classifiers contribute a feature that would otherwise be lacking in constructions with collectives, namely ‘individuation’. At the semantic level, this feature provides the access to individual elements of the counted set that is needed for the number assignment. At the syntactic level, we can account for it as a subcategorization requirement of the cardinal that needs to be satisfied in counting constructions. Plural numeral nominals fulfill this requirement through their number markers: given that nominal number markers provide a cardinal quantification, they need to bring with them the individuation necessary for this, hence a numeral TermP is already individuated and can enter a cardinal counting construction directly. In contrast to this, collective TermPs are not individuated yet, and hence require an external element to fulfill this function, namely a numeral classifier.

Numeral classifiers are usually related to a full noun, such as *head* in ‘four head of cattle’. In their usage as a classifier, though, they are semantically bleached when compared to this nominal base and restricted to a functional status: they provide the feature ‘individuation’, but—unlike measure words—do not contribute any lexical content.\(^\text{27}\) Thus, classifiers do not change the semantics of collectives, and neither do they map them from, say, substance-denotation to object-denotation: as discussed above, collectives, unlike mass nouns, have objects already as their denotation, that is, they refer to potentially countable entities.

The functional, semantically bleached status of classifiers is reflected in a number of characteristic phenomena cross-linguistically.\(^\text{28}\)

Their distribution is language-specific, rather than governed by general conceptual features, such that in different languages collectives with approximately the same meaning can be combined with different classifiers, whereas classifiers with similar nominal bases can be combined with different collectives. E.g., English *cattle* is combined with *head*, but its German counterpart *Vieh* is combined with *head*, changes the structure of the counting construction, even in these cases, it is the noun that determines whether this is a two-term (cardinal – noun) or three-term (cardinal – classifier – noun) counting construction.

\(^{27}\) Measure words are the linguistic elements that refer to measure units (cf. the discussion of measurement vs. counting in 4.2.1 above). Accordingly, they add their own lexical content to the phrase (= that of the measure unit), while classifiers do not contribute such additional information, but only the ‘individuation’ that provides access to individual elements of the counted set. As a result, nouns are associated with classifiers by way of subcategorization, but can be combined with measure units freely, only restricted by the measured property in question (e.g. it is only ‘four head of cattle’, not ‘four piece of cattle’ etc., but one can say ‘four litres of water’, ‘four kg of water’, ‘water of 4°C’, etc. with the respective differences in meaning). Cheng (this volume) distinguishes expressions for measure units as ‘massifiers’ from classifiers.

\(^{28}\) For detailed cross-linguistic overviews of classifier characteristics cf. Aikhenvald (2003); Grinevald (2000).
classifier Stück ‘piece’, while the German counterpart to head (Kopf) is a classifier for leafy vegetables that come in round shapes, such as salad and cabbage (e.g., ‘vier Kopf Salat’, literally ‘four head [of] salad’).

Counting constructions can involve something that looks like duplication at first sight, such as Burmese ‘cùñ hnà cùñ’, lit. ‘island two island’, where the same item, cùñ, occurs twice, once as a full noun contributing the lexical content (= the first occurrence), and once as a numeral classifier fulfilling the ‘individuation’ requirement (= the second occurrence of cùñ).

Numeral classifiers can be generalized and thus lose their proper classifying function, so that the same classifier can be used for all collectives (sometimes with a [- human] restriction), as for instance the Persian classifier tā (‘piece’ / ‘grain’). In German, the classifier Stück ‘piece’ serves as a default classifier for collectives that are not otherwise specified for a particular classifier.

Given that numeral classifiers often do not add any lexical content, they can sometimes remain implicit, making the construction similar to one with numeral nouns. In contrast to counting constructions with numeral nouns, though, collectives still appear in their bare form without number markers. Classifiers can either be optionally phonologically empty, e.g. in Persian or in Turkish, or generally implicit, e.g. in Hungarian. Again, this pattern is not restricted to languages with predominantly transnumeral nouns, but can also be found for collectives in plural languages, e.g. in English, where ‘four cattle’ can be found as well as ‘four head of cattle’.

Taken together, this, again, supports an account of collectives as a class of nouns that can be defined cross-linguistically, a class that behaves morphosyntactically like mass nouns, but denotes objects and thus has conceptual parallels with plural nouns, with the effect that collectives, just like plural nouns and unlike mass nouns, can appear in counting constructions. Due to the different status of their number markers, though, they cannot be pluralized in these constructions and involve an (implicit or explicit) functional element, a numeral classifier, to satisfy the cardinal’s ‘individuation’ requirement.

4.3 Variation in syntactic-conceptual mass/count correspondences

In sum, what we find in the morphosyntactic mass/count domain, then, is a distinction between transnumeral and numeral nouns that is related to the status of bare noun phrases and number marking and links up with the conceptual distinction of substances and objects in a many-to-many pattern, rather than in a strict one-to-one correspondence: while substance-denoting nouns (mass nouns) are transnumeral in principal, object-denoting nouns can be either transnumeral

(collectives) or numeral (plural nouns). The distinction between collectives and plural nouns, then, is independent of the kind of entities a nominal refers to: if a noun refers to an object, it can either be transnumeral or numeral, i.e. it can either obligatorily mark the distinction of one vs. many morphosyntactically, or not. Accordingly, we can find both options realized in the same language, as the example of collectives—and the corresponding classifier constructions—alongside plural nouns in English shows, and we can even find both numeral and transnumeral number marking realized in the same language, as the example of lexical singular marking for collectives in Bavarian shows.

Further support for an autonomy of syntactic and conceptual features comes from variation in syntactic-conceptual mass/count correspondences that highlights the linguistic arbitrariness realized in the distinction between collectives and plural nouns. In the present section, I illustrate this with examples from dialectal and diachronic variation, and from variation for the morphosyntactic reflexes of semantic coercions.

4.3.1 Dialectal and diachronic variation

In view of the conceptual similarity of collectives and plural nouns as object-denoting nouns, it is not surprising that we do not only find variation in the choice of one or the other option for corresponding nouns across languages, but also within languages, notably dialectal variation as well as diachronic changes from plural to collective behaviour and vice versa. (12) through (14) provide examples for this. (12) illustrates usages of nouns like cattle, furniture, or silverware as plural nouns, rather than collectives, which seems to be possible in some dialects (possibly including contact varieties of English, given that these are examples from internet forums). (13) and (14) illustrate diachronic changes.31 (13) provides data from Old Persian that indicates that at that stage, object-denoting nouns behaved as numeral nouns, appearing in two-term cardinal counting constructions in their plural form, rather than as collectives (= transnumeral), as in modern Persian.32 (14) contributes examples from Middle High German that point to a development in the opposite direction, namely from collectives to plural nouns (in New High German, the corresponding nouns would either be pluralized or combined with an indefinite article, just as in the English paraphrases, making the Middle High German text ambiguous from the point of view of translation into modern German):

31 For a detailed discussion of the development in Persian and German, see Wiese (1997).

32 From the beginning of the third century BC, there was a general tendency within the Iranian languages from synthetic to analytical NPs, whereby nouns lost inflectional endings. In accordance with this tendency, since Middle Persian, object-denoting nouns have moved from plural behaviour to predominantly collective behaviour.
(12) a. another two cattles ready for branding  
    [movie discussion]  
  b. I got the same four furnitures  
    [video game discussion]  
  c. How can we clean silverwares?  
    [household chemistry discussion]  

(13) a. Viyaxnahya māhyā ʁauca biš əkata əha  
    [Old Persian]  
    [name]_gen month_gen 14 days_pl gone were  
    ‘14 days of the month Viyaxnahya were gone. / It was the 15th day of  
    Viyaxnahya.’  

b. x̱ayaθiyə əgarbəyam  
    [Old Persian]  
    9 kings_acc,pl captured_i sg  
    ‘I captured nine kings.’  

(14) a. mäc hilfet wol, friunt verre baz  
    [Middle High German]  
    relative helps well friend much better  
    ‘{A relative is / relatives are} good help, {a friend / friends} a  
    much better one.’  

b. dä stüende ouch niemer ritters becher lære  
    [Middle High German]  
    there stood also never knight’s mug empty  
    ‘{A knight’s tankard / knights’ tankards} would never be empty there.’  

This synchronic (dialectal) and diachronic variation underscores the autonomy of  
conceptual and morphosyntactic distinctions in the mass/count domain. It shows  
that the morphosyntactic behaviour of nouns, and thus their classification as  
collectives or plural nouns, does not differ only across languages, but also can change  
within a language without affecting meaning, such that the same noun can be used as  
transnumeral or numeral, due to dialectal variation, or that whole nominal classes  
can change from numeral to transnumeral behaviour or vice versa, due to diachronic  
developments.

4.3.2 Variation in semantic coercions

Another class of phenomena reflects a kind of variation between numeral and  
transnumeral behaviour that prima facie seems to point in the opposite direction,

35 Subject of a posting: http://answers.yahoo.com/question/index?qid=1006051211863, last viewing Oct  
27, 2009.  
36 Data from DB (= Darius, inscription of Behistan) I.37f, cf. Kent (1953:81).  
37 Data from DB II.5-7, cf. Brandenstein and Mayrhofer (1964:85).  
namely to a close correlation of morphosyntactic and conceptual distinctions: the domain of semantic coercions. In the nominal mass/count domain, we find three major kinds of coercions, illustrated in (15) through (17). In each case (a) gives the basic usage, and (b) the coerced usage:

(15) **Grinder Coercion:**
- a. There {is a chicken / are chickens} in the yard. [numeral]
- b. There is chicken in the soup. [transnumeral]

(16) **Sorter Coercion:**
- a. She drinks tea. [transnumeral]
- b. The best teas are from India. [numeral]

(17) **Portion Coercion** ('Restaurant Talk'):
- a. She drinks tea. [transnumeral]
- b. We'll have two teas, please. [numeral]

As the examples illustrate, these coercions involve conceptual shifts yielding an enriched interpretation for the coerced nominals in the (b) cases: In grinder coercions, nouns that usually refer to an animal (15a) are used to denote a substance containing edible parts of this animal (15b), and in sorter and portion coercions, nouns that usually refer to a substance, e.g. a beverage (16a, 17a), are used to denote abstract objects, namely sorts (16b) or portions (17b) of this substance. That is, on the conceptual level we find shifts from objects to substances, or vice versa.

What makes this a potential challenge to our account of mass/count distinctions is that in each case, this conceptual shift is accompanied by a morphosyntactic shift: reference to objects is reflected by numeral behaviour of the nouns, reference to substances is reflected by transnumeral behaviour. Hence these coercions draw on a close one-to-one correlation of syntactic and conceptual distinctions, a correlation that supports the interpretation of basic vs. coerced usages. However, this support is not a necessary condition for mass/count coercions: a closer look at this domain shows that we also find coercions with nouns that remain in the transnumeral realm, both in their basic and in their coerced usage, making the shift one between mass nouns and collectives, rather than between mass and plural nouns. (18) and (19) give examples for the three kinds of coercions in Kurdish (Sorānī), where, like in modern Persian, nouns are generally transnumeral.

---

41 Cf. also Frisson and Frazier (2005) for processing evidence for such an enrichment, from an eye-tracking study with mass/count coercions.
42 The fact that there is no morphosyntactic reflex of coercions makes a nominal like masī 'fish' in (18) potentially ambiguous. Accordingly, the food interpretation is often supported by a specification like gošt-i 'meat of' (e.g. 'sordn-aka bē (gošt-i) merišk-a', 'The food is without (the meat of) chicken'). However, (18b) illustrates, grinder coercions can also work without such a specification. On restrictions on such coercions in Mandarin that are related to similar phenomena, cf. Cheng et al. (2008), Cheng (this volume).
Grinder Coercion:

a. masī.m kṛi [object(s): transnumeral noun]
   fish-1sg.erg bought
   ‘I bought {a fish / fishes}.’

b. xordn.aka bē masī.a [substance: transnumeral noun]
   eat-def without fish-is
   ‘The food is without fish.’

Sorter and Portion Coercion:

a. čāy axom [substance: transnumeral noun]
   tea drink.1sc
   ‘I am drinking tea.’

b. bašterīn čāy čāy.i hindīya [objects: transnumeral noun]
   best tea tea-of India.3sg
   ‘The best teas are from India.’

c. du (tā) čāy.mān bō bēna [objects : transnumeral noun]
   two cl tea-us for bring
   ‘Bring us two teas.’

Hence, in the domain of mass/count coercions as well, we find an autonomy of
syntactic and conceptual features that supports similar options across languages,
indipendently of syntactic number marking: we can coerce nouns from object
reference to substance reference and vice versa, in languages with numeral count
nouns as well as in those without, indicating that this coercion is independent of the
morphosyntactic status of nouns as numeral or transnumeral. Further evidence for
this comes from portion coercions in German restaurant talk, where nouns can also
remain transnumeral and behave as collectives, rather than plural nouns, when
referring to portions—just as in Kurdish, cf. (20):43

Portion Coercion ('Restaurant Talk'):

a. Ich trinke {Tee/Bier} [substance: transnumeral nouns]
   I drink {tea/beer}
   ‘I am drinking tea/beer.’

b. Zwei {Tee/Bier}, bitte. [objects: transnumeral nouns]
   two {tea/beer} please
   ‘Two teas/beers, please.’

The interesting—and rather surprising—aspect here is that German, unlike Kurdish, is a plural language, that is, object-denoting nouns behave as plural nouns as a rule, and collectives are an exception. Accordingly, German sorter coercions work just like English ones, such that nouns that are usually transnumeral (mass nouns) referring to substances become numeral (plural nouns) when referring to sorts, making sorter coercions with these nouns different from portion coercions:

Sorter Coercion

[objects: numeral nouns]

d. Die besten {Tees/Biere} kommen aus Indien.
the best {teas/beers} come from India

‘The best teas/beers come from India.’

Thus we have regular plural forms for these nouns, and these plural forms are systematically used in sorter coercions, but nevertheless, when it comes to portion coercions, nouns can remain transnumeral, pointing to a strong case of linguistic arbitrariness in the connection of syntactic and conceptual distinctions. To make things even more arbitrary, not all German nouns behave like that in restaurant talk: some do pluralize (e.g. Martinis), and others seem to be able to realize both options (e.g. one finds transnumeral Schnaps as well as plural Schnäpse), suggesting that this behaviour is subject to lexical specifications and intralexical variation.

This, then, further supports a view of a mediated, rather than direct correlation of syntactic and conceptual mass/count distinctions: the connection of the two levels varies in such a way that we find different options realized in different languages, within one language, and even for individual nouns. Given this freedom, the choice of particular patterns can then serve different purposes. A one-to-one correlation along the lines of ‘transnumeral→substance’, as we find it in English coercions, supports the interpretation by identifying basic vs. coerced usages of a noun, while the one-to-many correlation in German coercions can distinguish different kinds of coercions (sorts vs. portions).

4.4 Semantics as a mediator of syntactic and conceptual classifications

A level that has been suggested for the mediation of syntactic-conceptual correspondences, is that of grammatical semantics: if we distinguish semantics as part of the grammatical system from an extragrammatical, general conceptual system, then semantic representations can combine conceptual and syntactic distinctions in a way that accounts for one-to-one as well as many-to-many correlations in inter-

44 A mirror image of this might be Ojibwe: as Mathieu (this volume) shows, pluralized mass nouns are restricted to a portion reading, while for sorter interpretations, mass nouns are not pluralized but rather combined with a ‘kindifier’ that seems to work like a classifier for kinds.

and intralinguistic variation. (21) provides semantic representations for the nominal mass/count domain that capture this:46

(21) a. Mass nominals (beef): \( \exists x \ (\text{BEEF}(x)) \)

b. Collectives (cattle): \( \exists u \ \forall x \ (x \in u \rightarrow \text{COW}(x)) \)

c. Plural nominals (cows)

- "(a cow)" \( \exists u \ \exists V \ (\forall x \ (x \in u \rightarrow \text{COW}(x)) \land |V(u)| > 1) \)
- "(cows)" \( \exists u \ \exists V \ (\forall x \ (x \in u \rightarrow \text{COW}(x)) \land |V(u)| = 1) \)

In these representations the epsilon-operator, \( \varepsilon \), is used for the representation of indefinite terms, 'V' stands for an individuation function, and \(|V(u)|\) is the cardinality of an individuated set 'u'.47 The representations differ in two main respects, which leads to two overlapping classes that link up morphosyntactic and conceptual distinctions. (1) The semantic representations for collectives and plural nouns, but not those for mass nouns, have an internal structure that identifies individual elements: they involve a set 'u' consisting of elements 'x' that satisfy a certain predicate (e.g. COW). This accounts for the 'built-in modes [...] of dividing their reference' (Quine 1960:91) that such nouns possess, relating to the conceptual distinction of object-denoting vs. substance-denoting. (2) The semantic representations for plural nouns, but not those for collectives and mass nouns, involve an individuation function 'V' that provides access to individual elements of the set 'u'. This accounts for the different kinds of cardinal counting constructions that plural nouns and collectives appear in, and relates to the morphosyntactic distinction of numeral vs. transnumeral nominals.

Figure 4.3 illustrates how these semantic representations thus bring together syntactic and conceptual mass/count distinctions.

Hence, while on both the syntactic and the conceptual level we find two main nominal classes in the mass/count domain, the semantic level provides a three-fold distinction, with collectives in the intersection of conceptual vs. morphosyntactic mass/count classifications.

Processing evidence for such a three-fold distinction comes from a priming study we conducted with pairs consisting either of two collectives (= related pairs), or of a collective and a member of one of the other two classes, that is, a mass noun or a plural noun (= unrelated pairs), e.g. furniture-cattle (collective-collective) contrasted with furniture-beef (collective-mass) and furniture-cows

46 This is not the only possible semantic account of nouns and number marking, of course. What is important for our discussion (see below) is, on the one hand, the distinction between semantic representations that provide an internal structure and those that do not, and, on the other hand, the distinction between representations that contain an individuation function and those that do not. For an overview of the semantic representation of nominal number cf. Link (1991).

47 The epsilon operator is defined as: \( \exists x \ (Fx \land Gx) \iff \exists x(F(Gx)) \). On individuation functions in semantic representations, cf. for instance, Krifka (1995).
Reaction times for lexical decisions on the targets showed that collectives were processed significantly faster in the related pairs, that is, in the presence of another collective, compared to the unrelated pairs. We found this for English, German, and Spanish, that is, in plural languages where collectives constitute only a minor, exceptional nominal class, as well as for Persian, that is, in a transnumeral language where collectives are the rule for object-denotation. This indicates a psychological reality for collectivity as a cross-linguistic phenomenon: it indicates an activation of semantic collectivity that distinguishes collectives as a separate class in the nominal mass/count domain and holds across plural and transnumeral languages.

4.5 Conclusions

This chapter has been dedicated to a discussion of collectives, a nominal class that plays only a minor role in plural languages like English, but a major one in transnumeral languages like Mandarin or Persian, and indicates some central characteristics of the nominal mass/count domain: (1) The cross-linguistic occurrence of collectives shows that object-denoting nouns have basically two options for their morphosyntactic behaviour—they can be transnumeral (= collectives) or

---

48 Cf. Wiese and Piñango (2001); Piñango and Wiese (2004). We used a priming paradigm with lexical decision tasks; processing advantages on targets were measured by reaction times to the lexical decision. Targets were presented visually, primes were presented visually (uni-modal design) or auditorily (cross-modal design).

49 Note that in the Persian study, there were only collectives and mass nouns, but no plural nouns (given that object-denoting nouns are generally collectives, rather than plural nouns, in Persian).
The similarity of collectives and plural nouns at the conceptual level shows that the choice of these options is independent of a nominal’s interpretation, emphasizing an autonomy of syntactic and conceptual distinctions, rather than a one-to-one correlation. This autonomy supports a cross-linguistic mass/count distinction of nominals that is independent of the availability of syntactic plural marking in a language. The independence of morphosyntactic choices is further reflected in intralingual (diachronic and dialectal) variation between collective vs. plural behaviour of the same nouns, and by the range of different options realized across and within languages for morphosyntactic reflexes of mass/count coercions. From the point of view of linguistic architecture, the different ways to correlate morphosyntactic and conceptual distinctions can be captured by semantic representations that mediate morphosyntactic and conceptual distinctions and define three nominal classes in the mass/count domain, with collectives in the intersection of mass nouns and plurals.

Hence, collectives, despite the marginal role they might play in languages like English, form a legitimate, distinctive cross-linguistic class situated at the intersection of plural and mass nouns, one that highlights an independence of syntactic and conceptual classifications that bears not only on our view of mass/count distinctions and their relation to nominal number, but also on the way we model these distinctions within a linguistic architecture.
Individuation and inverse number marking in Dagaare

SCOTT GRIMM

5.1 Introduction*

Dagaare (Gur; Niger-Congo) exhibits an initially surprising system for marking number. The basic paradigm is given with respect to the Dagaare words ‘child’ and ‘seed’ in (1), showing a near\(^1\) minimal pair where both nouns share the same stem, yet the morpheme \(-\text{ri}^2\) marks the plural interpretation for ‘child’ and the singular interpretation for ‘seed’.

<table>
<thead>
<tr>
<th>Gloss</th>
<th>Singular</th>
<th>Plural</th>
<th>Stem</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1)</td>
<td>‘child’</td>
<td>bìé</td>
<td>bìrí</td>
</tr>
<tr>
<td>‘seed’</td>
<td>bìrì</td>
<td>bìè</td>
<td>bì-</td>
</tr>
</tbody>
</table>

Upon first view, this pattern would appear arbitrary and unstable in comparison with number marking systems from Indo-European languages—for how would one

\* This material is in part based upon work supported by the National Science Foundation under Grant No. 1023289. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of the National Science Foundation. This paper would not have been possible without generous efforts of the Dagaare scholars Mark Ali, Adams Bodomo, and J.A. Saanchi. In particular, Mark Ali devoted much time to discussing Dagaare vocabulary with me during visits to Ghana in 2008 and 2011. Dr. Bodomo has provided encouragement and aid at every step of the project. I would also like to thank the following for discussion and comments on various drafts: Matt Adams, Arto Anttila, Nadine Borchardt, Eve Clark, Jeff Good, Paul Kiparsky, Beth Levin, Uli Sauerland, Tom Wasow, and two anonymous reviewers who provided substantial comments resulting in many clarifications and improvements. I would also like to thank the organizers and audience of the Mass/Count Workshop held at the University of Toronto. This material has also benefited from being presented at Sinn und Bedeutung 14, California Universities Semantics and Pragmatics 1 at UCLA, Zentrum für Allgemeine Sprachwissenschaft, and the Department of African Studies at Humboldt University, Berlin and I would like to thank the audiences for much helpful feedback. The field research on Dagaare in 2008 was made possible by the generous support of the Stanford Center of African Studies. The usual disclaimers apply.

1 The pairs differ in tone and in that the plural of ‘child’ has a vowel which appears to have undergone lengthening. Anttila and Bodomo (2009) propose that Dagaare phonology makes use of lexically conditioned high vowel lengthening in such cases.

2 As marked in the examples, Dagaare possesses two levels of tone, high and low. The morpheme \(-\text{ri}\) itself is, however, unmarked for tone.
know if a particular noun is to be marked in the singular or the plural except on a noun-by-noun basis? Inverse number systems are rare but attested at least in North America (Kiowa) and the Pacific (Nehan).³ To give an example in a mock version of English, the pattern would appear as in (2), where -s is an inverse number marker in English:

<table>
<thead>
<tr>
<th>Gloss</th>
<th>Singular</th>
<th>Plural</th>
</tr>
</thead>
<tbody>
<tr>
<td>‘child’</td>
<td>child</td>
<td>child-s</td>
</tr>
<tr>
<td>‘seed’</td>
<td>seed-s</td>
<td>seed</td>
</tr>
</tbody>
</table>

This chapter, based on detailed investigation of the lexicon of Dagaare, argues that the core of Dagaare’s number marking system is based on the subtle but pervasive semantic principle of individuation, the propensity for an entity to appear as an individual unit. While individuation has most often been invoked in a binary fashion to account for the different behaviors of count and mass terms, I will show that it is also relevant when restricted to the count domain, where, for the case at hand, Dagaare’s morphology appears to be sensitive to different degrees of individuation. These results are relevant for how individuation is viewed in general. Some researchers have made a sharp distinction between individuated and non-individuated (Mufwene 1984, Bloom 1999), although others have viewed individuation as comprised of different sub-properties (Wierzbicka 1988) or as a continuum (Gentner and Boroditsky 2001). The view emerging from Dagaare’s morphology is compatible with the broader view of individuation where it does not simply distinguish in a binary fashion between count and mass nouns, but operates in a scalar fashion. Under this view, mass terms (water) are less individuated than entities typically occurring in collections (ribs), which are in turn less individuated than typically singular entities (dog). As such, individuation has potential effects throughout a given grammatical number system. This influence extends to the categorization of terms as collectives, common cross-linguistically for e.g. insects and vegetables/grains/fruit, as well as influencing the preference for occurring in the plural or singular, the distinction which will be shown to be crucial for the paradigm in (1). A broader understanding of individuation then provides a common semantic backdrop against which distinct patterns of grammatical number marking can be related, and of which the categories mass and count are simply the most visible endpoints of the scale.

A second theme throughout the chapter will be examining the Dagaare system from the perspective of markedness, as the observed morphological patterns are unexpected upon the standard view. Usually, the singular is considered unmarked and the plural is considered marked (see Greenberg 1966), an alignment which is clearly contradicted by the inverse number marking pattern. The following sections

³ For discussion of Kiowa, see Watkins (1980) and Harbour (2008). Nehan is discussed both in Corbett (2000) and Baerman (2007), both of whom provide general discussions of number marking reversals.
will also demonstrate that the data from Dagaare is consistent with markedness patterns; however, the universal tendencies are more nuanced than one would expect under the simple alignment between singular/unmarked and plural/marked.

This chapter is structured as follows. Section 5.2 discusses some of the basic facts of Dagaare’s grammatical number system and then proceeds to establish and test the primary hypothesis of this chapter that the distribution of -ri correlates with different levels of individuation: nouns unmarked in the singular pattern with highly individuated entities whereas nouns unmarked in the plural pattern with entities which are less individuated and/or tend to appear in groups. This general view is supported both by phenomena within Dagaare, such as the domain of diminutives and dialect variation as discussed in 5.3, and cross-linguistic facts that surface in a wide array of language types, discussed in 5.4. The chapter concludes with a formal analysis of the number marker -ri, demonstrating that this non-standard method of marking singular and plural distinctions can be aligned with standard semantic theories of number.

5.2 The semantic basis of inverse number marking in Dagaare

Before presenting evidence for a semantic basis of Dagaare’s nominal system, I lay out some of the basic features of Dagaare’s nominal system and discuss previous approaches. Additional examples of nouns similar to those in (1) are given in Table 5.1, where nouns marked by -ri in the plural are shown on the left and those marked by -ri in the singular are shown on the right. Table 5.1 also displays instances where -ri undergoes assimilation following nasals and liquids, resulting in the allomorphs -nl and -Il, respectively.

The forms in (1) and Table 5.1 can be shown to correspond to singular and plural interpretations by examining their interaction with other elements of the grammar that

<table>
<thead>
<tr>
<th>-V Singular</th>
<th>-rl^4/-nl Plural</th>
<th>Gloss</th>
<th>rl/-nl Singular</th>
<th>-V Plural</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>tìč</td>
<td>tìrì</td>
<td>‘tree’</td>
<td>lúgrí</td>
<td>lúgó</td>
<td>‘prop, pillar’</td>
</tr>
<tr>
<td>gbìé</td>
<td>gbèrì</td>
<td>‘forehead’</td>
<td>nyågrí</td>
<td>nyågá</td>
<td>‘root’</td>
</tr>
<tr>
<td>pič</td>
<td>pèrì</td>
<td>‘basket’</td>
<td>filì</td>
<td>filè</td>
<td>‘sore’</td>
</tr>
<tr>
<td>nànjà</td>
<td>nànnì</td>
<td>‘scorpion’</td>
<td>ìlì</td>
<td>ìllè</td>
<td>‘horn’</td>
</tr>
</tbody>
</table>

^4Capital letters for vowels indicate +/− ATR.

Vowels in Dagaare distinguish whether advance tongue root (ATR) is present or not. The standard orthography for Dagaare collapses the representations for the vowels /o/ and /u/ as well as for /e/ and /i/, but I have followed Bodomo (1997)’s orthographic system in which these distinctions are made. For the case at hand, this permits clearly seeing that -ri and -rî are allomorphs which differ only due to ATR-harmony.
mark number. Examples (3–4) demonstrate distinct number agreement with the singular (ŋâ) and plural (-má) forms of demonstrative pronouns. In the plural, the agreement prefix further distinguishes between human (ba-) and non-human (a-) referents.

(3) a. bíé ŋâ
ci últ.  DEM.PROX.SG
‘this child’
b. bíí-rí bà-má
ci últ. HUM.PL-DEM.PROX.PL
‘these children’

(4) a. bí-rì ŋâ
ci últ.  DEM.PROX.SG
‘this seed’
b. bíè à-má
ci últ. NHUM.PL-DEM.PROX.PL
‘these seeds’

Number words whose value is greater than one also show selection of plural forms and agreement. No agreement is visible, however, in the use of the word designating ‘one’, yéni, which acts as a modifier of the noun, directly attached to the noun stem. Number words designating ‘two’ or more select for plural nouns and again take plural agreement prefixes. Examples of both nouns from (1) combined with the number words for ‘one’ and ‘two’ are given in (5–6).

(5) a. bí-yéni (bi- + yéni)
ci últ. child-one
‘one child’
b. bíí-rí bá-yí
ci últ. HUM.PL-two
‘two children’

(6) a. bí-yénì (bi- + yéni)
ci últ. seed-one
‘one seed’
b. bíè à-yí
ci últ. NHUM.PL-two
‘two seeds’

Much previous work on the nominal system of Dagaare has considered the pattern in (1) from the perspective of a system of noun classes in Dagaare. Bodomo (1997), Kropp Dakubu (2005) and Bodomo and Marfo (2006) elaborate systems of

5 In Dagaare, modifiers such as adjectives are compounded with the noun stem and the adjective then supplies its own singular/plural suffix patterns.
nouns for Dagaare based upon different singular-plural pairings of nouns. The inverse marking pattern in these analyses is simply related to a distinction between two different singular/plural pairings: one is comprised of nouns ending in vowels in the singular and -ri in the plural, while the other is comprised of nouns ending in -ri in the singular and vowels in the plural.

Analyses of the different noun classes in Dagaare are clearly valuable from a diachronic and comparative perspective. As a Gur language, the nominal system of Dagaare stems from Proto-Gur, which possessed a highly developed noun class system (Miehe and Winkelmann 2007) and without a doubt, the pattern observed in (1) historically derives from a noun class system. At the same time, accounting for the inverse marking pattern is worthwhile in its own right. If the pattern in (1) were only a minor singular/plural pairing among many others, it would probably not merit much attention. Yet, in contrast to related languages, such as Gurenɛ (Nsoh 2002), which conserve more of the Proto-Gur system, the noun class system in modern Dagaare is largely decayed. The inverse pattern of (1) has become the predominant pattern of the nominal system, accounting for over 70 percent of the nouns in my current database which have singular and plural forms. The two other major singular/plural pairings are for nouns designating humans and for liquids and other typically uncountable nouns which possess a distinct plural, sometimes known as the ‘second plural’ in the Dagaare literature. Thus, accounting for the inverse marking pattern will lead to understanding the principle form of nominal organization in Dagaare.

A different perspective on singular and plural formation in Dagaare is provided by Anttila and Bodomo (2009), who provide detailed morphophonological analyses of the Dagaare nominal system. They uncover a range of regularities governing the morphophonology of Dagaare number inflection, which the account presented here is compatible with.

Building upon this previous research, this chapter considers the further question of what predictive factors, if any, govern the occurrence of nouns with -ri in the singular or -ri in the plural. Number marking in the nominal system of Dagaare is, from all appearances, not directly predictable from the phonological form of the stem. Counter-examples are provided by sets of minimal pairs, similar to the examples in (1) and Table 5.1, given in Table 5.2.

I now turn to demonstrating that the answer to this question can be found by examining the meaning of the nouns at issue.

5.2.1 Number marking and individuation

In English as well as cross-linguistically, nouns which refer to count entities differ from nouns which refer to mass entities in their morphosyntactic realization.

---

6 This classification is not a gender system in the sense of Corbett (1991) or Güldemann (2000), where genders are established based on agreement classes. Dagaare has little concord phenomena, and the only three agreement classes that can be established are those seen in examples (3)–(4): singular, human plural and non-human plural.
Commonly cited properties include the ability to be modified by certain quantifiers—mass nouns accept quantifiers such as ‘much’ or ‘little’ (‘much/little wine’) while count nouns accept quantifiers such as ‘many’ and ‘few’ (‘many/few books’), yet the converse does not hold for either class (‘?much books’/’?few water’). The literature attempting to account for such distinctions is vast and various, yet much of it reacts in one way or another to the principle of individuation. There are of course divergent perspectives on what individuation designates, but generally the thesis relates cognitive or perceptual qualities of objects to the grammatical realization of count and mass nouns. An early view from Quine (1960) held that count syntax provided an apparatus for individuating objects, viz. delimiting the relevant object from others and tracking its spatio-temporal identity, while mass syntax does not. This view leads to positing a sort of correspondence between syntax and entities in the world. On a strong version of this correspondence theory, language users should conceptualize the referents of count nouns as distinct, countable, individuated things and those of mass nouns as non-distinct, uncountable, unindividuated things (Wisniewski et al. 1996, p. 271). Varieties of this distinction have been picked up in the formal semantics literature, e.g. the atomic/non-atomic distinction in Link (1983), as well as in the psycholinguistic literature (e.g. Bloom 1994, Wisniewski, Imai, and Casey 1996, Barner and Snedecker 2005).

While the individuation hypothesis was primarily elaborated in relation to the mass/count divide, it is reasonable to suppose its influence is relevant within the count domain. First, although the divide between count and mass domains is often loosely spoken of as dichotomous, much work following on Allan (1980) has shown that not all countable nouns are created equal. Rather, evidence from interaction between different determiners and quantifiers demonstrates that there are different levels of countability between true count terms and uncountable mass terms. Accordingly, it is plausible that individuation is related to different levels of countability, and in turn, to the nominal morphology of Dagaare.

While individuation is a commonly cited concept, it suffers in the same manner as other commonly cited conceptual factors in linguistics, such as animacy and agenticity, in that individuation is far from rigorously defined. A rigorous definition will not emerge here either, but rather the strategy is to use individuation as a heuristic to

<table>
<thead>
<tr>
<th>Gloss</th>
<th>Stem</th>
<th>Singular</th>
<th>Plural</th>
</tr>
</thead>
<tbody>
<tr>
<td>‘wild rat’</td>
<td>kù-</td>
<td>kúó</td>
<td>kúúrí</td>
</tr>
<tr>
<td>‘hoe’</td>
<td>kù-</td>
<td>kùùrí</td>
<td>kùé</td>
</tr>
<tr>
<td>‘granary’</td>
<td>bùg-</td>
<td>bùgó</td>
<td>bùgrí</td>
</tr>
<tr>
<td>‘pillar’</td>
<td>lùg-</td>
<td>lúgrí</td>
<td>lúgó</td>
</tr>
</tbody>
</table>

Commonly cited properties include the ability to be modified by certain quantifiers—mass nouns accept quantifiers such as ‘much’ or ‘little’ (‘much/little wine’) while count nouns accept quantifiers such as ‘many’ and ‘few’ (‘many/few books’), yet the converse does not hold for either class (‘?much books’/’?few water’). The literature attempting to account for such distinctions is vast and various, yet much of it reacts in one way or another to the principle of individuation. There are of course divergent perspectives on what individuation designates, but generally the thesis relates cognitive or perceptual qualities of objects to the grammatical realization of count and mass nouns. An early view from Quine (1960) held that count syntax provided an apparatus for individuating objects, viz. delimiting the relevant object from others and tracking its spatio-temporal identity, while mass syntax does not. This view leads to positing a sort of correspondence between syntax and entities in the world. On a strong version of this correspondence theory, language users should conceptualize the referents of count nouns as distinct, countable, individuated things and those of mass nouns as non-distinct, uncountable, unindividuated things (Wisniewski et al. 1996, p. 271). Varieties of this distinction have been picked up in the formal semantics literature, e.g. the atomic/non-atomic distinction in Link (1983), as well as in the psycholinguistic literature (e.g. Bloom 1994, Wisniewski, Imai, and Casey 1996, Barner and Snedecker 2005).

While the individuation hypothesis was primarily elaborated in relation to the mass/count divide, it is reasonable to suppose its influence is relevant within the count domain. First, although the divide between count and mass domains is often loosely spoken of as dichotomous, much work following on Allan (1980) has shown that not all countable nouns are created equal. Rather, evidence from interaction between different determiners and quantifiers demonstrates that there are different levels of countability between true count terms and uncountable mass terms. Accordingly, it is plausible that individuation is related to different levels of countability, and in turn, to the nominal morphology of Dagaare.

While individuation is a commonly cited concept, it suffers in the same manner as other commonly cited conceptual factors in linguistics, such as animacy and agenticity, in that individuation is far from rigorously defined. A rigorous definition will not emerge here either, but rather the strategy is to use individuation as a heuristic to
gain insight into the nominal structure of Dagaare and consequently into the functioning of inverse number marking. I will consider four factors linked to the individuation hypothesis and their potential influence on the realization of nominals in Dagaare, which I now discuss along with the evidence for considering them relevant.

The first factor, animacy, receives a good deal of independent support. Animacy, relative to some sort of animacy scale ranging from humans to larger then smaller animals which in turn correlates to a scale of individuation, is known to influence number marking cross-linguistically (Smith-Stark 1974, Corbett 1996, 2000). In particular, the higher the entity designated by a noun rates on an animacy hierarchy, i.e. the closer to human, the greater the likelihood is that the noun expresses a singular/plural contrast.

Two other factors relating to individuation were proposed by Wierzbicka (1988), and subsequently investigated experimentally by Middleton et al. (2004). First, Wierzbicka argues that nouns designating entities for which the constituents are more easily distinguishable are more likely to be used as count nouns, while those nouns designating entities for which the constituents are not easily distinguishable will be used as mass nouns. For example, she argues that beans is more likely to be a count term than rice since individual beans are in principle easier to distinguish than individual grains of rice. Middleton et al. (2004) examined this hypothesis experimentally, where subjects had to match a nonce count or mass term with one of two graphical displays of novel aggregates which varied in distinguishability. The graphical displays of novel aggregates were sets of 40 elements where ‘each element was a simple shape with a black-to-white gradient that appeared slightly 3-dimensional and did not obviously resemble the constituents of any familiar aggregate’ (p. 382). They then presented subjects with pairs of aggregate displays which varied along the dimensions of spatial proximity to other elements (Close versus Apart) and size of elements (Large versus Small). Thus, a subject would see two sets of an element where for one set, each element was spatially separated from the other and for the other set each element was spatially contiguous with other elements. The subject would then decide which picture aligned with a phrase such as ‘This is worgel’. The general results were that subjects’ choices of count or mass terms were significantly influenced (p < .001) by spatial proximity, but not by the size, of the elements. These results are compelling as the design of the experiment using nonce items ensures that such factors are general.

The second factor argued for by Wierzbicka (1988) is the canonical manner of interaction with a given entity. She exemplifies this with examples such as the naming of berries in Polish, which are generally count terms because, she claims, people interact with them one by one, viz. picking/eating them, while farmers selling berries typically use mass syntax to describe berries since they interact with them in
quantities rather than individually. This factor was investigated via novel objects, again by Middleton et al. (2004). They presented subjects with a novel aggregate, ‘yellow decorative coarse-grained sugar’, in a cardboard box, which the subjects then needed to match to one of two phrases presented in count and mass syntax (e.g. ‘This is worgel/These are worgels’). The experimenters manipulated the mode of interaction with the aggregate. In the baseline condition, the subjects simply observed the material and then were presented with a response sheet to decide which phrase was appropriate. In the interaction condition, the experimenter and the participants used a thin paper-clip implement to scoop up individual grains of the material and insert each grain into a hole of a board distinct from the box containing the material. The participants then were presented with the response sheet to decide which of two phrases was appropriate, one with mass and one with count syntax. The responses for the baseline and interaction conditions were inversely related: a majority of participants in the baseline condition (69 percent) selected a mass phrase while a majority of participants in the interaction condition (61 percent) selected a count phrase. While this result is not definitive, it would appear that the mode of interaction with an aggregate can affect the manner by which it is referred to.

The final factor I consider is the likelihood of a noun to be ‘inherently plural’, in other words the likelihood that individual referents of a noun canonically appear as a member of a pair or group, as, for example, is the case for paired body parts (e.g. kidneys). Recent work by Acquaviva (2008) has emphasized the distinctive morpho-semantic behavior of entities which canonically appear in collectives, duals and other ‘marked’ number categories. While individuation is normally considered only in light of mass/count syntax, it seems probable that entities that canonically appear as a member of a pair or group, as in the case of duals and collectives, are qualitatively different from those which canonically appear as individuals. This distinction is independent from the previous factors. With regard to animacy, Corbett (1996) previously pointed out that dual/collective paradigms are orthogonal to the animacy scale. Very general number marking patterns, such as the occurrence of plural marking, are correlated with the animacy scale: the higher on the animacy scale the referent of a noun falls, the more likely it is to allow plural marking. Yet, nouns which accept or require dual/collective marking do not systematically align with the animacy scale, thus, such a factor is independent of animacy. Further, the factors of distinguishability and interaction were examined by Middleton et al. (2004) by use of stimuli that only presented groups of entities, which in effect held constant whether a particular type of entity was more likely to appear within a group or singly, again making it plausible that the factor of inherent plurality is distinct from distinguishability or interaction.

7 This distribution is significantly above chance (p<.05).
5.2.2 Individuation and inverse marking

The individuation factors discussed immediately above have been argued, primarily in the case of English, to independently affect the realization of number marking. If individuation is grammatically relevant, a clear hypothesis emerges in considering the inverse number system of Dagaare: the more likely the entity is to be viewed as individuated, the more likely the singular noun will be unmarked and -\( ri \) will mark the plural; and conversely, the more likely the entity is to be viewed as coming in groups or non-individuated, the more likely the plural noun will be unmarked and -\( ri \) will mark the singular. The information about a noun’s individuation level therefore would be lexical information. Nouns would come with a basic number, determined by its semantic properties, while application of -\( ri \) gives the inverse value. This can be schematically pictured as in (7):

\[
\text{(7)} \quad \begin{array}{c}
\text{[Highly Individuated } N]\ \to +\ -ri \Rightarrow \text{ plural} \\
\text{[Less Individuated/Inherently Plural } N]\ \to +\ -ri \Rightarrow \text{ singular}
\end{array}
\]

The considerations of individuation lead straightforwardly to testable predictions. If individuation has an effect on the distribution of -\( ri \), one should observe distributional asymmetries in the appropriate semantic domains. In part guided by how transparent verification in a lexicon would be, I explored four relevant predictions:

(i) Nouns for higher-level (more salient) animals are more likely to be unmarked in the singular than nouns for insects (animacy).
(ii) Nouns for trees should be unmarked in the singular in comparison to nouns for vegetation (distinguishability).
(iii) Nouns for tools should be more likely to be unmarked in the singular than the converse (one canonically interacts with them individually).
(iv) Nouns for body parts which inherently come in pairs or groups should be more likely to be unmarked in the plural than not; while nouns for body parts which inherently come in single units should be more likely to be unmarked in the singular than not (inherently plural).

I now turn to the results of fieldwork which bear on these hypotheses.

5.2.3 Results from fieldwork

To test the predictions elaborated in the preceding section, I conducted fieldwork in Ghana with native speakers to develop a wordlist to determine the behavior of inverse number marking. The findings below are based on a wordlist of nearly 1500 words which I compiled during my field research.\(^8\)

\(^8\) I would like to acknowledge Arto Anttila and Adams Bodomo for generously permitting me to incorporate elements from their wordlist, Anttila and Bodomo (2006).
As the hypotheses involved distributions over semantic domains, I coded each word for (relatively transparent) semantic domains, where possible. The chart in Figure 5.1 displays the results with respect to hypotheses (i–iii). The x-axis displays various semantic domains while the y-axis displays the number of lexicon entries. The dark-shaded regions show the number of lexicon entries in a given semantic domain with the singular unmarked, while the light-shaded regions show the number which are unmarked in the plural and marked by -\textit{ri} in the singular. For instance, the category of \textit{mammal} shows 43 entries in the lexicon that are unmarked in the singular and 5 entries which are unmarked in the plural and marked in the singular by -\textit{ri}. In these counts, I excluded derived forms, since they follow their own patterns which tends to obscure any generalization. Nominal derivation will be discussed in section 5.2.4.

Figure 5.1 demonstrates reliable asymmetries visible across the semantic domains. Nouns for higher-level animates, namely mammals, birds, and reptiles are typically unmarked in the singular; however, the majority of nouns for insects have a plural that is unmarked. Similarly, nouns for trees are typically unmarked in the singular, while most nouns for vegetation are unmarked in the plural. Nouns for tools, which were hypothesized to be individuated as a result of the typical manner with which one interacts with them, also showed a strong tendency towards being unmarked in the singular.

Figure 5.2 shows similar results for the fourth hypothesis, viz. nouns for body parts which inherently come in pairs or groups should be more likely to be unmarked in the plural, while nouns for body parts that inherently come as singular items should be more likely to be unmarked in the singular. The x-axis displays

![Figure 5.1 Number marking across semantic domains](image-url)
whether the noun is inherently singular, e.g. the term for *head* where canonically humans only have one, or inherently dual/plural, e.g. *eye* or *rib* where canonically humans have two or multiple of each, respectively. Again the $y$-axis displays the number of items in the lexicon for each category.

5.2.4 Discussion

The above results indicate that Dagaare morphology is sensitive to the degree of individuation for the referent of a noun, i.e. -ri marks the singular when a noun is considered to be low in individuation/inherently plural, otherwise it marks the plural. Section 5.2.1 laid out a number of hypotheses which made specific predictions about particular semantic domains. When the individuation hypothesis is applied systematically to the lexicon of Dagaare, it uncovers many exceptions; however, most of these demonstrate semantic sub-regularities, or result from practices elsewhere in the grammar, viz. derivational morphology, or historical considerations. I next discuss each of these in turn.

5.2.4.1 Semantic sub-regularities  
Nouns that do not conform to the general trend of the domain often display semantic sub-regularities. One instance from the animate domain is that most of the nouns for insects unmarked in the singular are for insects capable of causing harm (e.g. scorpion, wasp, spider). In the domain of tools, while the vast majority of nouns were marked by -ri, or one of its allomorphs, in the plural, exceptions included nouns such as fúmini (sg) / fúminé (pl) ‘needle’ or mìri (sg) / mìè (pl) ‘rope/string’, both of which are consistent with the individuation hypothesis. While ‘needle’ would appear a strong candidate for canonically appearing in a collection, the canonical rope for Dagaare speakers
consists of two strands twisted together, which again would be consonant with inherent plurality. (Cf. English ‘twine’ which contains the root for ‘two’.)

5.2.4.2 Derived forms Several types of derivational processes are in action in Dagaare, which cloud the homogeneity of any semantic generalizations. In particular, the general number marking trend of a semantic domain may not be realized by a particular member of the domain if its singular-plural pattern is simply due to its derivational history.

Nouns formed by compounding take the number pattern associated with the final element of the compound. While in the domain of nouns for insects, singulars are typically marked by -ri, three instances are given in Table 5.3 where the noun is formed by compounding, and the noun’s number marking follows that of the final member of the compound.

Table 5.3. Compound forms

<table>
<thead>
<tr>
<th>Singular</th>
<th>Plural</th>
<th>Gloss</th>
<th>Derivational Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>bōndōnāá</td>
<td>bōndōnni</td>
<td>‘insect’ (general)</td>
<td>bōn ‘thing’ (n.) + dōn ‘bite’ (v.)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>+ -aa (agentive suffix)</td>
</tr>
<tr>
<td>kyānkyāmpōnāà</td>
<td>kyānkyāmpōnni</td>
<td>‘firefly’</td>
<td>kyānkyān ‘regularly’ (adv.)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>+ pūn ‘flash’ (v.)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>+ -aa (agentive suffix)</td>
</tr>
<tr>
<td>sēŋkpōglí</td>
<td>sēŋkpōglò</td>
<td>‘bedbug’</td>
<td>sēŋ ‘bed’ (n.) + kpōglī ‘lump’ (n.)</td>
</tr>
</tbody>
</table>

Similarly, exceptions arise in the domain of humans. Nouns for humans normally fall into a distinct singular-plural pairing, taking the plural marker -ba; however, some do not, especially those that designate human roles (‘doctor’, ‘lord’, ‘strong person’, etc.). Most of these items exhibit marking by -ri for the plural, as one would expect in the human domain under the individuation hypothesis. Yet, there is again a set of examples for which the singular is marked by -ri, or an assimilated form thereof, a sampling of which is provided in Table 5.4. These examples all are related to verbs, from which they are ostensibly derived.9 Thus, derivational history can again account for the semantic misalignment of these lexical items under the individuation hypothesis.

A final source of misalignment between the present semantic hypothesis and the lexical items of Dagaare is provided by their etymology.

5.2.4.3 Borrowing Number marking for words borrowed into Dagaare appears to be influenced by phonetic similarity, as can be seen from the example lō̌̀rī (sg) lṑè

9 Dagaare seems to possess a semi-regular process whereby nouns can be zero-derived from verbs and adjectives, although this subject needs further research.
(pl) ‘truck; lorry’, clearly a borrowing from the English word *lorry*. The Dagaare singular form *lɔ̀rì* corresponds to the English singular form *lorry*, and the Dagaare plural is formed on analogy with words for which the stem is vowel final and the singular is marked by *-ri*. A similar explanation can be found for the term *tìrí* (sg) / *tì* (pl), borrowed from Akan (Mark Ali, p.c.). Table 5.5 shows three Dagaare words which display the pattern which is extended to *lɔ̀rì* and *tìrí* by analogy.

### Table 5.4. Humans marked in the singular by *-ri*

<table>
<thead>
<tr>
<th>Singular</th>
<th>Plural</th>
<th>Gloss</th>
<th>Derivational Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>gbèrì</td>
<td>gbèè</td>
<td>‘cripple’</td>
<td>gbèrì(v.) ‘to cripple’</td>
</tr>
<tr>
<td>sènì</td>
<td>sènìmé</td>
<td>‘girlfriend’</td>
<td>sèη(v.) ‘to befriend’</td>
</tr>
<tr>
<td>zùlì</td>
<td>zùlò</td>
<td>‘fool’</td>
<td>zùlì(v.) ‘to be foolish’</td>
</tr>
</tbody>
</table>

#### 5.2.4.4 Semantic shift

Some of the lexical entries which mark their singular with *-ri* have primary meanings that appear to be at odds with the individuation hypothesis, yet they retain traces of an antecedent meaning which accords with the hypothesis. An instance of this phenomenon is *yírì* (sg) / *yíè* (pl) ‘house’. The physical entity ‘house’ would appear to be a canonical example of an individuated entity. But, as often is the case, the canonical gloss obscures the full range of meanings associated with the word. While synchronically the most common use of *yírì* is to designate ‘house’, it has antecedent meanings as ‘compound’ (Durand 1953) as well as ‘family’ or ‘family members’ (Mark Ali, p.c.). In its uses for ‘compound’ and ‘family’, *yírì* designates something closer to a collection of entities, i.e. units of the compound/family members, and thus is more aligned with the notion of inherently plural/collective entities.

### Table 5.5. Borrowing and paradigmatic analogy

<table>
<thead>
<tr>
<th>Stem</th>
<th>Singular</th>
<th>Plural</th>
<th>Gloss</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>gbè-</td>
<td>gbèrì</td>
<td>gbèè</td>
<td>‘leg’</td>
<td></td>
</tr>
<tr>
<td>le-</td>
<td>lèrì</td>
<td>lèè</td>
<td>‘bead’</td>
<td></td>
</tr>
<tr>
<td>lo-</td>
<td>lòrì</td>
<td>lòé</td>
<td>‘meteorite’</td>
<td></td>
</tr>
<tr>
<td>loo-</td>
<td>lòrì</td>
<td>lòè</td>
<td>‘truck; lorry’</td>
<td>English</td>
</tr>
<tr>
<td>ttr-</td>
<td>tìrí</td>
<td>tìè</td>
<td>‘spoon’</td>
<td>Akan</td>
</tr>
</tbody>
</table>
Dagaare is sensitive to degrees of individuation and to collections as opposed to individuals. This section adduces further data from Dagaare that validates the implications of this assumption, and as such increases the plausibility that Dagaare is organized in such a fashion.

5.3.1 *Dialect variation*

I have argued that individuation plays an organizing role in the choice of nominal inflection in Dagaare. Given that degrees of individuation are akin to a scale-structure, one would expect to see dialect variation in the mid-region of the scale, i.e. entities which are not clearly individuated or group-like would be predicted to vary. Adams Bodomo (p.c.) has noted that there are instances where the direction of number marking differs among dialects, as shown for the noun stem *pi-* ‘rock’ in Table 5.6.

<table>
<thead>
<tr>
<th>Table 5.6. Dialect variation: Variation in directionality of marking</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Singular</strong></td>
</tr>
<tr>
<td>-----</td>
</tr>
<tr>
<td><em>piiri</em></td>
</tr>
<tr>
<td><em>pié</em></td>
</tr>
</tbody>
</table>

While nouns such as ‘human’ and ‘rib’ are naturally associated with individual and collective interpretations, respectively, items such as ‘rock’ could in principle be associated with either individual rocks or collections of rocks. Such claims must be assessed through further research, yet even so, aligning number formation with the propensity towards individuation provides an explanation for variation where purely morphophonological considerations would be hard-pressed to do so.

A second source of variation is found in the choice between *-ri* and a singulative marker, *-ruu*, which Dagaare employs to designate ‘a piece of’ for a limited set of nouns. The singulative appears mainly with clear mass terms as well as aggregates which are particularly close-knit, as shown in Table 5.7, where due to the nasal environment the singulative marker is realized as *-nuu*.

Several words in Dagaare, as shown in Table 5.8, vary across dialects between whether a noun marks its singular form with *-ri* or *-ruu*. This dialect variation in

<table>
<thead>
<tr>
<th>Table 5.7. Singulative paradigm</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Singular/Base</strong></td>
</tr>
<tr>
<td>-------------------</td>
</tr>
<tr>
<td><em>móó</em></td>
</tr>
<tr>
<td><em>súnní</em></td>
</tr>
</tbody>
</table>
Table 5.8. Dialect variation: Singular vs. singulative

<table>
<thead>
<tr>
<th>Variant</th>
<th>Singular</th>
<th>Plural</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>I:</td>
<td>váálí</td>
<td>váálá</td>
<td>‘rubbish’</td>
</tr>
<tr>
<td>II:</td>
<td>váálóó</td>
<td>váálá</td>
<td>‘rubbish’</td>
</tr>
<tr>
<td>I:</td>
<td>kómmírí</td>
<td>kómmié</td>
<td>‘tomato’</td>
</tr>
<tr>
<td>II:</td>
<td>kómmíríúú</td>
<td>kómmié</td>
<td>‘tomato’</td>
</tr>
</tbody>
</table>

Turn supports the main hypothesis that -ri marks the singular for objects which are inherently plural. The use of the two different markers implies that there is overlap between inherently plural and mass/aggregate terms.

5.3.2 Domains of the diminutive

Another area of Dagaare morphology where sensitivity to the collections/inherently plural distinction occurs is in the domains of its diminutive morphemes. Dagaare actually makes use of two separate diminutives, -lee (‘small’) and -biri (‘seed’) illustrated in Table 5.9, each of which highlights different associated meanings. In the instance of -lee, the application of the diminutive marker results in the meaning ‘small/young’, while the application of -biri derives lexical items that are associated with the base noun, often through a part/whole relation. Of particular interest is that -biri is used predominately for describing entities which come in collections, viz. toes, fingers, germs/bacteria (‘sick seeds’), beads, bullets, ribs, words (‘speech seeds’), stars, bees (‘honey seeds’). Uses of -biri occasionally describe items which do not come in collections, but which are associated with the base noun through relations

Table 5.9. Diminutives in Dagaare

<table>
<thead>
<tr>
<th>Diminutive</th>
<th>Stem</th>
<th>Gloss</th>
<th>Derived Noun (Sg)</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>-lee</td>
<td>bà</td>
<td>‘dog’</td>
<td>báléé</td>
<td>‘puppy’</td>
</tr>
<tr>
<td></td>
<td>náábò</td>
<td>‘cow’</td>
<td>nááléé</td>
<td>‘calf’</td>
</tr>
<tr>
<td></td>
<td>gànggáá</td>
<td>‘drum’</td>
<td>gànggáléé</td>
<td>‘alto drum’</td>
</tr>
<tr>
<td></td>
<td>gbé-</td>
<td>‘leg/foot’</td>
<td>gbébírí</td>
<td>‘toe’</td>
</tr>
<tr>
<td></td>
<td>nú-</td>
<td>‘hand’</td>
<td>núbírí</td>
<td>‘finger’</td>
</tr>
<tr>
<td></td>
<td>si</td>
<td>‘honey’</td>
<td>sibírí</td>
<td>‘bee’</td>
</tr>
<tr>
<td>-biri</td>
<td>málfâ-</td>
<td>‘gun’</td>
<td>málfâbírí</td>
<td>‘bullet’</td>
</tr>
<tr>
<td></td>
<td>yèl-</td>
<td>‘say, tell (v.)’</td>
<td>yêlbírí</td>
<td>‘word’</td>
</tr>
<tr>
<td></td>
<td>bàâl-</td>
<td>‘sick’</td>
<td>bàâlombírí</td>
<td>‘germ’</td>
</tr>
</tbody>
</table>
such as ‘small’ or ‘part of’, as well as through the literal meaning ‘seed of’, but these are comparatively rare.

Languages more often employ just one diminutive from which all the diminutive senses are derived (see the survey in Jurafsky 1996). An example is provided by Ewe, which, like Dagaare, has a diminutive which is taken from the word for ‘child’, -ví; however, -ví applies across both senses seen in Dagaare, resulting in derivations for ‘young/small X’ as well as items that come in collections (‘toes’), as seen in Table 5.10 (data from Heine and Kuteva 2009).

In contrast to the more common pattern, where one morpheme is generalized to many related senses, Dagaare marks the collective sense separately from other diminutive functions. In sum, -lee and -biri are not equivalent: -lee is closer to a standard diminutive generally meaning ‘young’ or ‘small’ while the employment of

<table>
<thead>
<tr>
<th>Stem</th>
<th>Gloss</th>
<th>Derived Noun (Sg)</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>nûtsu</td>
<td>‘man’</td>
<td>nûtsu-ví</td>
<td>‘boy’</td>
</tr>
<tr>
<td>yevú</td>
<td>‘European’</td>
<td>yevú-ví</td>
<td>‘young European’</td>
</tr>
<tr>
<td>nûi</td>
<td>‘cow’</td>
<td>nûi-ví</td>
<td>‘calf’</td>
</tr>
<tr>
<td>kpé</td>
<td>‘stone’</td>
<td>kpé-ví</td>
<td>‘small stone’</td>
</tr>
<tr>
<td>du</td>
<td>‘small village’</td>
<td>du-ví</td>
<td>‘small village’</td>
</tr>
<tr>
<td>afɔ</td>
<td>‘foot, leg’</td>
<td>afɔ-ví</td>
<td>‘toe’</td>
</tr>
</tbody>
</table>

-biri most often yields a collective sense, highlighting another area where Dagaare is sensitive to a distinction between individuated entities and collections.

### 5.4 Cross-linguistic correlates

Another source of support for the assumptions underlying the primary hypothesis that individuation underlies the organization of Dagaare’s nominal number marking system can be adduced from cross-linguistic data. Section 5.2 identified a set of semantic domains that are typically unmarked in the plural. In the same manner in which one expects certain features of the mass terms to be consistent across languages, viz. not accepting cardinal terms without a measure term, one would expect the behavior of the nouns associated with semantic domains unmarked in the plural to have parallel behavior across languages. In this section, I will show that cross-linguistic correlates to the unmarked plural in Dagaare surface in an array of language types, demonstrating that, despite different encodings, these systems seem to make similar divisions on a scale of individuation.

---

10 According to my current database, twenty nouns derived with -biri result in some sort of collective sense similar to those nouns shown in Table 5.9, six derived nouns indicate the literal ‘seed of’, which again result typically in a collection of entities, while the associated senses of ‘small’ and ‘part of’ occur only with two nouns each.
5.4.1 Nominal paradigms

Semantic domains similar to those discussed in section 5.2 are cross-linguistically relevant for collectives and duals (see discussion in Acquaviva 2008). One example that accords quite well with the findings in Dagaare is the collective/singulative class in Welsh, discussed in Stolz (2001). Welsh dispenses of a singular/plural distinction for count nouns just as in, say, English, where the plural is morphologically marked; however, for select semantic domains, a collective interpretation is morphologically unmarked, while a singular interpretation is signalled by a singulative marker -yn or -en. For instance, the Welsh term for ‘flea’ is chwain, a collective term whose singulative form is chwann-en. Of interest are the semantic domains where this holds: small animals and insects, vegetables/cereals/fruits, body parts (‘ribs’, etc.), and what Stolz terms ‘uncountable substance’, essentially granular mass terms (‘turf’, ‘embers’, ‘sand’), all of which accord with the semantic domains seen as unmarked plurals in Dagaare. A similar division is in effect for languages with nominal class systems, e.g. Swahili (Contini-Morava 2000) and Lingala (Mufwene 1980), where some noun classes appear to be unmarked in the plural. Once again, the relevant semantic domains are strikingly similar to those in play for Dagaare, e.g. vegetation, pairs and collectives.

5.4.2 Morphological behavior

The semantic domains discussed in section 5.2 also manifest unexpected behavior with respect to morphological processes. Tiersma (1982) noted that classes of nouns for entities that ‘naturally occur in pairs or groups’ tend to show surprising behavior with respect to morphological leveling, borrowing and double plural formation. Tiersma claims that ‘when a referent of a noun naturally occurs in pairs or groups, and/or when it is generally referred to collectively, such a noun is locally unmarked in the plural.’ Local markedness then is a semantic notion of markedness, which may effect how nouns are morphologically marked. This characterization aligns well with the distributional patterns established for Dagaare and converges with the individuation hypothesis. This section reviews the evidence from morphological processes which support the notion of ‘unmarked plurals’.

5.4.2.1 Morphological leveling Tiersma (1982) provides a detailed discussion of instances of paradigm leveling, a form of analogical change whereby a paradigm regularizes, as in the change from Early Modern English reach/raught to reach/reached. Tiersma notes that while morphological paradigms typically level towards the unmarked members of the paradigm, in certain cases nominal paradigms level in

---

11 Granular mass terms in Dagaare have several realizations, but most often only have a form marked by -ri, as in zəgi ‘flour’.
favor of the plural stem. This is exemplified in Modern Frisian which exhibits ‘breaking’ in conservative forms of singular/plural pairs, i.e. an alternation between falling and rising diphthongs, shown in Table 5.11 for the pairs /ie/ and /ji/ and /oa/ and /wa/. This alternation is undergoing leveling, where typically the falling diphthong, which is associated with a singular stem, is generalized; however, for a small number of items, the leveling generalized instead towards the rising diphthong, associated with the plural stem. As the selections of nouns in Table 5.11 indicate, this occurs in instances where the referent of the noun ‘naturally occurs in pairs or groups’.

<table>
<thead>
<tr>
<th>Conservative</th>
<th>Innovative</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>kies (sg)/ kjizzen (pl)</td>
<td>kjizze (sg) / kjizzen (pl)</td>
<td>‘tooth’</td>
</tr>
<tr>
<td>toarn (sg)/ twarnen (pl)</td>
<td>twarne (sg) / twarnen (pl)</td>
<td>‘thorn’</td>
</tr>
</tbody>
</table>

5.4.2.2 Borrowing As with leveling, borrowing typically proceeds by taking the unmarked singular stem; however, there are cases, and not surprisingly in the same semantic domains, in which the plural form is borrowed in preference to the singular. A clear example is provided by Welsh borrowings from English shown in Table 5.12, discussed in Stolz (2001), where the borrowed plural form from English serves as the basic term which can then be inflected for the singulative.

<table>
<thead>
<tr>
<th>Singulative</th>
<th>Collective</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>ffigys-en</td>
<td>ffigys</td>
<td>figs</td>
</tr>
<tr>
<td>gwsbery(s)-en</td>
<td>gwsberys</td>
<td>gooseberries</td>
</tr>
<tr>
<td>pys-en</td>
<td>pys</td>
<td>peas</td>
</tr>
</tbody>
</table>

5.4.2.3 Double plurals The formation of double plurals occurs when older plurals have been reanalyzed as singular units, leading to the addition of another plural marker. Tiersma (1982) provides examples from West Frisian which disposes of two productive plural markers, -en and -s. The examples in Table 5.13 show both markers have been applied. The referents here again involve an element of inherent plurality, which ostensibly provides a motivation for the reanalysis, as the plural form is the more basic, or ‘locally unmarked’ form.
Evidence from English frequency patterns

Throughout my discussion of Dagaare, and the above morphological patterns from other languages, I have referred to ‘unmarked plurals’. This term has been appropriate insomuch as these plurals have had less morphological material as opposed to a clearly suffixed singular, and thereby qualify as basic. Yet, for an element to be ‘unmarked’ has another implication in terms of text frequency, as in Greenberg (1966) where unmarked forms were shown to have greater text frequency than marked ones. If the semantic domains I have discussed are truly unmarked, one would expect to see ‘unmarked plurals’ in languages which do not display any morphological evidence of such a pattern reflected in terms of text frequency. In order to evaluate this prediction, I examined frequencies for the semantic domains of animal and insect from the COBUILD corpus (18 million words) provided by CELEX. Using basic terms and terms consistent with the vocabulary of Dagaare, I calculated the plural-to-singular ratio for these two domains, shown in Figure 5.3, where the x-axis represents the ratio of the token frequency of plurals to the token frequency of singulars and the y-axis represents the number of lexical items. The graph indicates that there is a clear trend for insect terms to have a plural/singular

<table>
<thead>
<tr>
<th>Singular</th>
<th>Plural</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>boei</td>
<td>boejens</td>
<td>‘handcuff’</td>
</tr>
<tr>
<td>lears</td>
<td>learzens</td>
<td>‘boot’</td>
</tr>
<tr>
<td>reed</td>
<td>redens</td>
<td>‘skate’</td>
</tr>
<tr>
<td>trep</td>
<td>treppens</td>
<td>‘stair’</td>
</tr>
</tbody>
</table>

Table 5.13. Double plurals in West Frisian (from Tiersma 1982)

5.4.3  Evidence from English frequency patterns

Figure 5.3 Number marking frequency patterns in English for Animals and Insects
ratio greater than 1, i.e. insect terms occur more frequently in the plural, while animal terms tend to have a plural/singular ratio less than 1, i.e. animal terms occur more frequently in the singular. This finding lends additional support to the arguments about morphological patterns in this section, as well as to the assumption that the plural of certain domains is ‘unmarked’.

5.5 A formal account of -\textit{ri}

I have established that Dagaare is sensitive to the degree of individuation and inherent plurality associated with the referents of nominal elements. Indeed, this should not come as a surprise since such facts have been consistently assumed with respect to the count/mass divide. Here, I have made the case that individuation is a matter of degree, sensitivity to which pervades the nominal system and is not limited to partitioning nouns into count and mass terms. This section will demonstrate that once inherent plurality, and thereby singularity, become available as lexical information, the semantics of inverse number marking can be aligned with formal models of the semantics of number. Further, the empirical data from Dagaare bears upon a controversy among different popular analyses of the plural.

Most analyses of the semantics of number since at least Link (1983) base their model of the count domain on two ingredients: (i) a set of atomic objects in some domain \( A \), which correspond to individual entities such as a dog or a chair and (ii) a domain \( E \), where \( E \subseteq A \), containing sets generated from the atomic objects in \( A \), which is structured by a part-whole or subset relation relating the sets of atoms. The singular of a noun denotes the atomic entities for which the noun is a true description, while the plural of a noun denotes, at least, sets of atomic entities for which the noun is a true description. The whole domain, \( E \subseteq A \), possesses the structure of a join semi-lattice. Beyond the basic agreement on these two ingredients, however, there has been substantial debate as to what is contained in the domain of the plural.\(^{12}\)

One line of proposals, originating with Link (1983), models the plural as denoting the closure of atoms under join (\(\oplus\)) less the atoms themselves, thus the denotation of the plural excludes that of the singular. In this treatment, the singular (atomic) denotation is simpler, and thus the unmarked form.

Many researchers, including Krifka (1989) and Sauerland et al. (2005), have noted that this account is problematic for environments such as negation or in questions. For instance, if the plural designates \textit{two or more}, then the statement \textit{Ed didn’t see horses} would be true if Ed had seen only one horse, but that is clearly not the desired result. Such facts have been taken to motivate a weaker plural designating \textit{one or more}, whereby the denotation of the plural includes that of the singular, in formal

\(^{12}\) See Farkas and de Swart (2010) for extensive discussion of the debate.
terms corresponding to the entire semi-lattice structure of the denotation of a noun. Under this analysis, the singular is more specific than the plural, and the plural surfaces as the unmarked number.

The two sides of this debate differ on whether (i) the singular or plural is unmarked and (ii) the model-theoretic structure of the domain of plurals—whether the plural is better modeled as intrinsically exclusive \((\text{sums} - \text{atoms})\) or inclusive \((\text{sums} \cup \text{atoms})\). As to the first point, the data from Dagaare reviewed above reveal a more intricate picture. Dagaare demonstrates that the cross-linguistic facts are more complicated than if only the singular or plural were unmarked, rather markedness is conditioned upon a nominal’s level of individuation. In order to evaluate the second point in light of the data from Dagaare, I consider two analyses, one consistent with the exclusive and the other consistent with the inclusive plural. Applying the logic of both exclusive and inclusive plural analyses to data from Dagaare shows the exclusive plural analysis has better empirical traction in inverse number marking systems.

5.5.1 The exclusive plural

Assuming inherent plurality—and thereby singularity—becomes available as lexical information—as argued in the above sections—and assuming the exclusive plural analysis, the semantics of inverse number marking is relatively straightforward: \(-ri\) is simply treated as a form of negation of the unmarked number value for the noun in question. This is an intuitive version of the function of inverse number marking, and is in essence a formal semantic update of the analysis of Kiowa in Wonderly (1954). Further assuming along with Ojeda (1998) that the base or root of the noun has a denotation of the entire space generated by the atoms and their sums \((\text{atoms} \cup \text{sums})\), i.e. the base is compatible with singular and plural individuals, then \(-ri\) can be modeled as the operation of complementation \((C)\), relativized to the domain of the base. The degree of individuation determines whether a noun is considered lexically plural or singular, whereupon \(-ri\) applied to a lexically singular noun will yield a plural denotation, while if \(-ri\) is applied to a lexically plural noun, it will yield a singular denotation.

Representative derivations are given in Table 5.14, demonstrating that this analysis clearly secures the desired interpretations. In prose, for lexically singular nouns, the application of \(-ri\) gives the complement of the denotation of a singular noun, viz. the complement of the relevant set of atoms. The value returned is the sums formed from the atoms, less the atoms themselves, which is in turn exactly the value of the noun’s plural denotation. For lexically plural nouns, the application of \(-ri\) gives the complement of the denotation of a plural noun, viz. the complement of the relevant set of sums. The value returned is the atoms which form the sums, which is in turn exactly the value of the noun’s singular denotation.

13 This line has also been developed independently in Bach (2007) and Bach (2008) for Kiowa.
The inclusive plural

An alternate analysis, which is consistent with weak plural analyses, models -ri as designating the completion of the space.\textsuperscript{14} The weak plural analysis of English plurals claims that the plural is unmarked, denoting closure under join, while the singular, designating atoms, is more specific. When the plural form is used, the singular interpretation is excluded by pragmatic blocking. The same inferences found in negative and interrogative contexts in English which motivate the weak plural analysis were also elicited in Dagaare, thus one could think to analyze -ri when marking the plural just as the English plural is analyzed, designating closure under join, with the singular interpretation disallowed by blocking.

The inclusive plural analysis can be extended to -ri by positing a parallel analysis for nouns which are lexically plural. In this case, the application of -ri to lexically plural nouns must also yield the entire semi-lattice, whereupon the plural interpretation is disallowed by blocking. This can be achieved through modeling -ri as forming the ideal of the denotation of the plural, the set of sums. A subset of a lattice is an ideal if it is a non-empty downward closed set\textsuperscript{15} which is also closed under join. When applied to the set of sums, closure under join is already satisfied while downward closure results in including all the atoms.

Ideal formation will also secure the desired result for lexically singular nouns, whereby -ri can be uniformly analyzed as the closure of the space under ideal formation. Downwards closure is already satisfied since by definition atoms have no proper parts, and closure under join results in the inclusion of all the sums. The singular interpretation then is disallowed by blocking. Representative derivations are given in Table 5.15, where $\mathcal{C}$ represents the ideal formation closure operator. In prose, for lexically singular nouns, the application of -ri gives the closure of the

\begin{table}
\centering
\caption{Derivations of lexically singular (‘child’) and plural (‘seed’) nouns with the exclusive plural}
\begin{tabular}{|c|c|}
\hline
Lexically Singular & Lexically Plural \\
\hline
\texttt{[[bi-] := \lambda x(CHILD(x))} & \texttt{[[bi-] := \lambda x((SEED(x))^{\odot} - SEED(x))} \\
\texttt{[[bi-] + ri} & \texttt{[[bi-] + ri} \\
\texttt{([\| bi- \|])}^{\mathcal{C}} & \texttt{([\| bi- \|])^{\mathcal{C}}} \\
\texttt{[\lambda x(CHILD(x))]^{\mathcal{C}}} & \texttt{[\lambda x((SEED(x))^{\odot} - SEED(x))]^{\mathcal{C}}} \\
\texttt{\lambda x((CHILD(x))^{\odot} - CHILD(x)} & \texttt{\lambda x[SEED(x)]} \\
\texttt{= PL(bi-)} & \texttt{= SG(bi-)} \\
\hline
\end{tabular}
\end{table}

5.5.2 The inclusive plural

\textsuperscript{14} I am indebted to Uli Sauerland for suggesting this line of analysis.

\textsuperscript{15} A formal definition of downward closure in a lattice is if $L$ is a lattice and $J$ is a non-empty subset, then $a \in L$, $b \in J$ and $a \leq b$ imply $a \in J$ (Davey and Priestley 2002, p. 44).
denotation of a singular noun, which is the entire semi-lattice. For lexically plural nouns, the application of -ri gives the closure of the denotation of a plural noun, which is again the entire semi-lattice.

Behavior under negation would demonstrate whether such a suggestion was feasible, for in many languages, such as English and as was elicited in Dagaare, negation of the plural always also excludes the truth of the singular. Extending the inclusive plural analysis to -ri in this way predicts that the negation of the form marked by -ri should exclude the truth of both singular and plural. The example in (8) shows that this turns out not to be the case (Adams Bodomo p.c.):

(8) n dà bá dà bè (zàà)
1st.pro.sg neg buy pst seed.pl (any)
I didn’t buy (any) seeds.

The same entailment patterns hold in Dagaare as in English and in (8) the negated plural also indicates that the speaker did not buy a single seed; however, the form negated is not marked by -ri, but rather is the unmarked form. As the data does not align with the logic of the inclusive plural analysis, this analysis must be rejected in favor of the exclusive plural analysis.

In concluding this section, it is noteworthy that the ability to align the number system of Dagaare with standard semantic accounts indicates that the semantics of number in Dagaare are similar to better described languages, in other words Dagaare speakers may communicate similar meanings, it is just the means of expression, here the morphology, that differs.

5.6 Conclusion

This chapter has demonstrated that number marking in Dagaare, and more generally inverse number marking, while at first sight surprising, under closer inspection can be seen to be a clever exploitation of relatively universal markedness

<table>
<thead>
<tr>
<th>Table 5.15. Derivations of lexically singular (‘child’) and plural (‘seed’) nouns with the inclusive plural</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lexically Singular</td>
</tr>
<tr>
<td>[bi-] := λx(CHILD(x))</td>
</tr>
<tr>
<td>[bi-] + ri (ippo [bi-])[^1]</td>
</tr>
<tr>
<td>[λx(CHILD(x))][^2]</td>
</tr>
<tr>
<td>λx((CHILD(x))[^3])</td>
</tr>
<tr>
<td>= inclusive(bi-)</td>
</tr>
</tbody>
</table>
patterns: less individuated/inherently plural entities which are semantically un-marked in the plural are also morphologically unmarked in the plural. Once the lexical generalizations are clear, the formal implementation is straightforward. Such systems reinforce the point that number marking comprises more than simply marking a dichotomy between reference to atoms or reference to sums, rather the canonical properties of the referents themselves influence how the marking is achieved. Articulating the precise connections between the lexical properties of nouns and number realization remains a fertile area for further exploration.
General number and the structure of DP*

ILEANA PAUL

6.1 Introduction

The goal of this chapter is to investigate the structure of DP by looking at one particular phenomenon: general number (Corbett 2000). In particular, I will argue that general number correlates neither with a lack of functional structure in the DP, nor with obligatory narrow scope (contra e.g. Rullmann and You 2006). As such, this chapter can be seen as part of the growing body of literature on the connections between the syntactic structure of DP and semantic interpretation. I will mainly draw on data from Malagasy (Western Austronesian), but I will also discuss data from other languages.

General number is a phenomenon that has been described by Corbett (2000: 9–19) as nominal forms that are interpreted as entailing ‘one or more x’. In languages with general number, the meaning of a noun can be expressed without any reference to number. Corbett (2000: 11), citing Dick Hayward (p.c.), provides an example of general number in Bayso (Cushitic).

(1) lúban foofe. [Bayso]
    lion.general watched.1.sg
    ‘I watched lion.’ (it could be one, or more than that)

* I would like to thank the Malagasy consultants who helped me with the data: Rita Hanitramalala, Jean Christophe Jaonesy, Tsiorimalala Randriambololona, Vololona Rasolofoson, Francine Razafinboaka, Martelline Razafindravola, and Rado Razanajatovo. I would also like to thank Sandy Chung, Lisa Matthewson, Hotze Rullmann, audiences at UBC, at the Mass/Count workshop at the University of Toronto, and at AFLA XVI at UC Santa Cruz, and two anonymous reviewers for their comments. Data are from my own notes, unless otherwise indicated. All errors are my responsibility. This research was supported by the Canada Research Chair program, SSHRC (SRG410-2005-1758), and the University of Western Ontario.
General number can be manifested in different ways in different languages; for example, in Bayso, it is also possible to overtly mark the noun as singular, paucal or plural. In fact, it seems likely that all languages have some means to indicate number, whether on the noun itself or within the noun phrase (but see Everett 2005, 2009 for claims that Pirahã has no number marking at all). One of the goals of this chapter is to examine the different manifestations of general number and what they tell us about DP structure and interpretation. In other words, does general number correlate with other syntactic and semantic properties of DPs?

This chapter is organized as follows. I first provide some data on Malagasy nouns in section 6.2. Given that general number is typically associated with bare nouns, I will focus on the distribution and interpretation of bare nouns. For the purposes of this chapter, ‘bare noun’ is used for any nominal that lacks a determiner. The presence/absence of an adjective or other modifiers does not appear to play a significant role in Malagasy. Section 6.3 explores general number and bare nouns in other languages and analyses that have been proposed. In section 6.4, I argue that despite claims in the literature, there is no one-to-one correlation between the syntax and semantics of DPs, that the mapping between morphosyntax and semantics is not rigidly determined (see Chung 2000 for similar conclusions).

6.2 Malagasy

Malagasy is a language with fairly rigid VOS word order. Bare noun arguments—nominals with no determiner—are possible (but only in non-subject position). Thus while (2a,b) are grammatical, (2c) is not.

(2) a. Manolotra penina izy.
   at.offer pen 3(nom)
   ‘She offers a pen/pens.’

   b. Rakofana kopy ny tsaramaso.
   tt.cover cup det bean
   ‘The beans are covered with a cup/cups.’

   c. Tonga mpanatra.
   arrive student
   ‘A student arrived.’

Although bare nouns are not permitted in subject position, they are otherwise found in a wide range of syntactic positions, such as direct object (e.g. (2a,b)), predicate, object of a preposition, and possessor.

Malagasy has a dedicated pre-nominal definite determiner ny (unmarked for number), but there is no indefinite determiner. In fact, the only number marking
The examples in (3) illustrate a singular and plural demonstrative, respectively.

(3) a. Omeo ahy itsy boky itsy.
   give.IMP 1SG(ACC) DEM book DEM
   ‘Give me that book.’
   (Rajemisa-Raolison 1971: 54)

b. Omeo ahy iretsy boky iretsy.
   give.IMP 1SG(ACC) DEM.PL book DEM.PL
   ‘Give me those books.’

Pronouns are also marked for number, but the third person pronouns only optionally show number with the addition of a plural demonstrative, as illustrated in Table 6.1.

Now that we have seen the basic distribution of bare nouns, I turn to their interpretation.

6.2.1 General number

I will first discuss general number as it is manifested in Malagasy. It has long been noted that nouns in Malagasy are unmarked for number—so (2a,b) can be translated in different ways (Keenan 1976). The most accurate, though awkward, translation is ‘one or more’ (e.g. ‘She offers one or more pens’). I first show that Malagasy nouns are indeterminate for number and not ambiguous. I then turn to some particular interpretive properties of nouns that have been noted to correlate with general number. I should point out here that in the data presented below, I have used bare nouns (nouns without a determiner), but the same results hold for DPs (headed by an overt determiner). This fact will be relevant later in the discussion of cross-linguistic variation.

6.2.1.1 Ambiguity In order to test whether nouns are ambiguous between singular and plural or simply indeterminate, I use one of the standard tests. Consider first a

1 It is of course possible to create a plural reading with the addition of a numeral or quantifiers.
case of ‘true’ ambiguity: the verb mamaky means either ‘to cut’ or ‘to read’. If mamaky is used in the following context (4), we can see that once the meaning has been set in the first clause, the ‘do so’ clause can only share that reading.2

(4) Namaky i Soa; toraka izany koa i Be.
pst.at.read/cut det Soa same that too det Be
‘Soa read and so did Be.’ or ‘Soa cut and so did Be.’
not ‘Soa read and Be cut.’ or ‘Soa cut and Be read.’

Thus mamaky patterns with other truly ambiguous words such as bank in English.

Turning now to nouns, we see a different pattern emerge. In (5), the precise number of books is not at issue. Therefore Soa could have read one book and Be two. The value of singular versus plural remains undetermined.

(5) Namaky boky i Soa; toraka izany koa i Be.
pst.at.read book det Soa same that too det Be
‘Soa read one or more books; so did Be.’

In other words, in (5) both Soa and Be read an indeterminate number of books; the number of books that they each read can be different. This result shows that bare nouns are indeterminate for number rather than ambiguous, as is typical of general number (see e.g. Bliss 2003, Rullmann and You 2006, Sato 2008).

6.2.1.2 Discourse anaphora

Given that nouns are unmarked for number, we expect certain syntactic effects in the grammar. In particular, we expect bare nouns to be the antecedents to either singular or plural pronouns. As seen below, the bare noun namana ‘friend’ can be the antecedent to a singular or a plural pronoun, as in (6).3

(6) a. Namangy namana aho omaly; nanasa
pst.at.visit friend 1sg(nom) yesterday pst.at.invite
azy (ireo) hiara- misakafo.
3(acc) dem.pl together at.eat
‘I visited one or more friends yesterday; I invited him/them to dinner.’

b. Namaky boky aho ary Rakoto
pst.at.read book 1sg(nom) and Rakoto
no nomeko an’ izy io / ireo.
foc pst.tt.give acc 3 dem / dem.pl
‘I read one or more books and it was Rakoto that I gave it/them to.’

---

2 The determiner i in the following examples is for certain proper names.
3 Note that as we saw in Table 6.1 the pronoun azy is itself unmarked for singular or plural.
The examples in (7) also illustrate general number: the bare noun vorona ‘bird’ is unmarked for number and is therefore compatible with either continuation, but a noun modified by iray ‘one’ is only compatible with disjunction.

   pst.at.see bird 1sg(nom) duck or pigeon / duck and pigeon  
   ‘I saw one or more birds—a duck or a pigeon / a duck and a pigeon.’

b. Nahita vorona iray aho – ganagana na fony / #ganagana sy fony.  
   pst.at.see bird one 1sg(nom) duck or pigeon / duck and pigeon  
   ‘I saw one bird—a duck or a pigeon / a duck and a pigeon.’

The facts in (6) and (7) fall out directly if nouns in Malagasy are unmarked for number and pattern with similar facts in Mandarin and Catalan, as discussed by Rullmann and You (2006) and Espinal (2010), respectively.

6.2.1.3 Pragmatic effects Rullmann and You (2006) point out that nouns that show general number do not have scalar implicatures (unlike nouns modified by numbers). I show here that the same facts hold in Malagasy. Thus it is possible to follow up the first clause in (8a), which contains an unmodified noun, with a clause that makes precise the number of books. The example in (8b) shows that such a precision with a noun marked for number is pragmatically odd.

   pst.at.buy book 1sg(nom) yesterday five foc pst.tt.buy.1sg(gen)  
   ‘I bought one or more books yesterday. I bought five.’

b. Nividy boky iray aho omaly. #Dimy no novidiko.  
   pst.at.buy book one 1sg(nom) yesterday five foc pst.tt.buy.1sg(gen)  
   ‘I bought one book yesterday. I bought five.’

Similar effects can be seen in (9).

(9) a. Manana fiara izy, angamba roa.  
   at.have car 3(nom) maybe two  
   ‘She has one or more cars, maybe two.’

b. Manana fiara iray izy, # angamba roa.  
   at.have car one 3(nom) maybe two  
   ‘She has one car, maybe two.’

The data in (10) indicate that metalinguistic negation of the scalar implicature ‘not more than one’ is possible with a noun modified by iray ‘one’ (10b), but not with unmodified nouns (10a).
(10)  a. Tsy nividy boky aho, #fa nividy dimy.
    neg pst.at.buy book 1sg(nom) c pst.at.buy five
    ‘I didn’t buy one or more books, I bought five.’

   b. Tsy nividy boky iray aho, fa nividy dimy.
    neg pst.at.buy book one 1sg(nom) c pst.at.buy five
    ‘I didn’t buy one book, I bought five.’

Summing up, what we have seen thus far is that nouns in Malagasy show general number and that this is a sign of indeterminacy rather than ambiguity that has both syntactic (anaphora) and pragmatic effects (implicatures).

6.2.2 Scope

I now look more specifically at bare nouns and their scopal properties. First, I note that bare nouns are indefinite and are therefore used to introduce novel referents into the discourse (see Paul 2009b for more discussion). As an illustrative example, consider (11).

(11) Nisy zazakely ao an-trano. #Nahafantatra zazalahy aho.
    pst.at.exist child there acc house pst.at.know boy 1sg(nom)
    ‘There were children in the house. I knew boys.’
    (consultant’s reaction: ‘the two sentences don’t go together’)

The first sentence introduces zazakely ‘children’ and it is not possible to refer back to a subset of that group with the bare noun zazalahy ‘boys’.

Once we recognize that Malagasy bare nouns are indefinite, their scopal properties become important. It has long been noted that bare nouns take obligatory narrow scope in a range of languages. For example, Carlson (1977) shows that English bare plurals always scope low with respect to other quantificational elements.

(12) Everyone read books on caterpillars. (∀>∃ but *∃>∀)

Similar facts have been documented for Mandarin (Rullmann and You 2006), Turkish (Bliss 2003), Indonesian (Chung 2000), Javanese (Sato 2008), and many other languages. In fact, some authors argue that bare nouns are incorporated and that incorporation leads to their special scope behaviour (see Farkas and de Swart 2003). Van Geenhoven (1999) refers to this process as ‘semantic incorporation’.

Turning now to the Malagasy data, we see that bare nouns take variable scope. The examples in (13) illustrate the scope of the bare noun dokotera ‘doctor’ with respect to the opaque verb te ‘want’. Example (13a) shows the narrow scope reading and (13b) illustrates the wide scope reading.

4 I am not probing the possibility of anaphora between the bare noun and a pronoun, per se—I am interested in the wide scope reading, which is made salient by using a pronoun in the above examples. As
(13) a. Te hanam- bady dokotera aho nefa . . .
   want FUT.AT.have spouse doctor 1SG(NOM) c
   ‘I want to marry a doctor but . . .’

   b. . . . mbola tsy mahita.
   still NEG AT.see
   ‘. . . I still haven’t found one.’

   c. . . . mipetraka lavitra ahy izy.
   AT.live far 1SG(ACC) 3(NOM)
   ‘. . . he lives far from me.’

The examples in (14) show both scope readings for the bare noun alika ‘dog’.

   AT.look-for dog 1SG(NOM) or dog what or dog what
   ‘I’m looking for a dog—any dog.’

   b. Mitady alika aho – kely sy mainty ilay izy.
   AT.look-for dog 1SG(NOM) small and black DEF 3(NOM)
   ‘I’m looking for a dog—it’s small and black.’

Similar facts obtain with modals, such as tokony ‘should’: (15a) illustrates narrow scope for boky ‘book’ and (15b) illustrates wide scope.

(15) a. Tokony hamaky boky ianao –
   should FUT.AT.read book 2SG(NOM) –
   na boky inona na boky inona.
   or book what or book what
   ‘You should read a book—any book.’

   b. Tokony hamaky boky ianao –
   should FUT.AT.read book 2SG(NOM) –
   ‘farihy manga’ ny anarany.
   lake blue DET name,3(GEN)
   ‘You should read a book—“Blue Lake” is its title.’

Finally, the data in (16) show how bare nouns interact with other quantificational elements, such as the universal quantifier rehetra ‘all’ in (16a) and negation in (16b,c).

   PST.AT.read book French DET student all
   ‘All the students read a French book.’ (∀>∃ or ∃>∀)

argued by many authors (e.g. Mithun 1984, Dayal 2007, Farkas and de Swart 2003), anaphora depends less on the nature of the bare/incorporated noun and more on the nature of pronouns in the particular language.
b. Tsy nahasitrana zaza ny dokotera. Marary loatra ilay izy.
   neg pst.cause.cure child det doctor sick too dem 3(nom)
   ‘The doctor was not able to cure a child. He (the child) was too sick.’

c. Tsy nanam- bady dokotera aho satria
   neg pst.at.have spouse doctor 1sg(nom) because
   nipetraka lavitra ahy izy.
   pst.at.live far 1sg(acc) 3(nom)
   ‘I didn’t marry a doctor, because he lived far from me.’

Summing up, Malagasy bare nouns take variable scope, a surprising property from a cross-linguistic perspective. The scope facts deserve much further discussion, but for the purposes of this chapter I simply note that the Malagasy data show that there is no direct correlation between bare nouns and narrow scope and therefore no correlation between general number and narrow scope. See Paul (2009a) for further discussion.

6.2.3 Determiner non-plurality effect

In his discussion of number in Singlish, Gil (2003) notes that although bare nouns show general number and determiners are compatible with singular or plural, the addition of a determiner forces a singular interpretation (in the absence of an overt numeral). Gil refers to this effect as the ‘Determiner non-plurality effect’. The relevant data are provided below (Gil 2003: 473–474).

(17) a. Geraint eat apple [Singlish]
           b. Geraint eat the apple
c. Geraint eat that apple
d. Geraint eat the apples
e. Geraint eat that apples

Example (17a) illustrates that Singlish has bare noun arguments and general number: apple may be interpreted as singular or plural (or mass). The subsequent examples show that the determiner and demonstrative are compatible both with singular and plural nouns. Crucially, however, in (17b,c) we have the non-plurality effect: rather than being unmarked, the resulting interpretation is singular (or mass). The plural reading is impossible to obtain. Gil remarks that this effect is found in a range of languages, including Japanese (Downing 1996), Mandarin and Singaporean Malay. He notes, however, that the effect is not present in Cantonese, Hokkien, Standard Malay or Riau Indonesian: the equivalent of (17b) is underspecified for number.

---

5 As discussed by Gil, the addition of a possessor also induces a non-plurality effect. Given that some of the languages Gil discusses do not have determiners, per se, the effect seems to be more generally connected with the addition of material in the function domain of DP.
To that list, we can now add Malagasy. Consider the following example, a slightly
modified version of (2b):

(18) Rakofana ny kopy ny tsaramaso.

\textit{The one or more beans are covered with the one or more cups.}

Both DPs \textit{ny kopy} ‘the cup(s)’ and \textit{ny tsaramaso} ‘the bean(s)’ show general number.
I have found many examples in texts where a DP can be either singular or plural—
the context can of course make one reading or the other more salient. For example,
in a story where the king has told the people to take their spears and throw them at
the intruder, it is clear that \textit{ny lefona} ‘the spear’ is plural.

(19) ...fa nitoraka ny lefona avy lavitra.

\textit{... and (the people) threw the spears from afar.} (Ravololomanga 1996: 38)

Thus there is no determiner non-plurality effect in Malagasy.

6.3 Cross-linguistic considerations

As noted earlier, there is a growing body of literature on the syntax and semantics of
bare nouns, literature that draws on data from a wide range of languages. The notion
of general number often arises in these discussions. In this section, I will consider
certain analyses of bare nouns and general number, and how Malagasy fits into the
overall pattern.

6.3.1 Chierchia 1998

Chierchia’s very influential paper (1998b) explores the distribution of bare noun
arguments from a cross-linguistic perspective. Chierchia correlates count nouns with
number inflection and predicts that bare noun arguments are mass, both within and
across languages. As mass nouns, they are not countable directly and therefore
require a classifier in the context of numerals. Chierchia’s basic semantic parameter
states that languages can vary in the lexical meaning of nouns. Lexical nouns can be
of type \(<e,t>\) or \(e\) and only the latter can occur bare in argument positions. Type \(e\)
nouns are kind terms and their basic denotation (according to Chierchia’s definition)
is mass. Lexical nouns of the predicative type \(<e,t>\) can be count or mass and
can be converted to an argumental type with the addition of a determiner (the
determiner may be \(\emptyset\), as in Italian, or the language may allow covert typeshifting, as
in Slavic). Chierchia posits two features that give rise to three classes of languages.
Languages where nouns are of type \(e\) are \([ + \text{arg}, -\text{pred}]\) (e.g. Chinese); languages
where nouns are of type \(<e,t>\) are \([-\text{arg}, +\text{pred}]\) (e.g. French); and languages where
nouns can be of either type are \([ + \text{arg}, +\text{pred}]\) (e.g. English).
Many criticisms of Chierchia’s proposal have appeared in the literature, but I will not undertake a thorough review here. Instead, I would like to point out that for Chierchia, there is a correlation between generalized bare noun arguments and mass denotation. Although Malagasy has generalized bare noun arguments, it is easy to show that there is also a robust grammatical mass-count distinction (and no classifier system). Thus the quantifier *tsirairay ‘each’ is compatible with count nouns, such as *boky ‘book’, but not with mass nouns, such as *lafarina ‘flour’.

(20) a. Novidiny ny boky tsirairay.  
   pst.tt.buy.3(gen) det book each  
   ‘She bought each book.’

b. *Novidiny ny lafarina tsirairay.  
   pst.tt.buy.3(gen) det flour each

Similarly, it is possible to have weak quantifiers as the predicate with count nouns (21a), but not mass nouns (21b).

(21) a. Roa/vitsy ny boky.  
   two/few det book  
   ‘There are two/few books.’

   two/few det flour

The Malagasy pattern fits with Wilhelm’s (2008) conclusion that the distribution of classifiers has less to do with the mass versus count denotation of nouns and more to do with the nature of the numerals themselves.

To make Malagasy fit within Chierchia’s system requires two steps. First, general number does not correlate with one particular setting of the parameter and can in principle be observed in any language (this move weakens the role that number inflection plays in Chierchia’s typology, but is not inconsistent with the overall proposal). Second, Malagasy is like Slavic, a [+arg, +pred] language, but with an overt definite determiner. If this is the correct setting of the parameter for Malagasy, we expect bare nouns to be used to refer to kinds. Testing for kind reference in Malagasy, however, turns out to be difficult. As mentioned above, bare nouns cannot appear in the subject position and I have yet to find a truly kind-selecting predicate. When asked to translate terms such as invent, speakers suggest mamorona, which has a more general meaning of ‘to form, to create, to produce’ and is not a kind-selecting predicate. Moreover, judgements about sentences with mamorona and a

---

6 It should be noted that for Chierchia, the correlation is not bi-directional. He predicts that if the nouns in a language denote in the mass domain, the language will have generalized bare noun arguments. There is no implication in the opposite direction (if a language has generalized bare noun arguments, the nouns in the language will denote in the mass domain); so for him, Malagasy is not a counterexample.

7 I have yet to find an element that selects for only mass nouns in Malagasy.
bare object are not clear-cut: some speakers allow a kind reading for the bare noun, but others do not.

Summing up, under Chierchia’s analysis, general number must be an epiphenomenon and does not give any direct clues into either the semantics or the morphosyntax of bare nouns.

6.3.2 Rullmann and You 2006
As far as I am aware, Rullmann and You (2006) is the only article to consider general number within the context of Chierchia’s analysis. They propose that general number is in fact the key notion to understanding Mandarin bare nouns, rather than the mass-count distinction. Crucially they argue that languages with generalized bare noun arguments show general number. Moreover, general number is shown to correlate with obligatory narrow scope in a wide range of languages. Rullmann and You do not provide a formal analysis of the connection between narrow scope and general number, but they suggest that bare nouns lack certain functional layers. And perhaps the lack of structure leads to low scope.

There is a long list of languages that can be cited in support of this correlation between bare nouns, general number and low scope (e.g. Mandarin, Turkish (Bliss 2003), Indonesian (Chung 2000), Javanese (Sato 2008), Hebrew (Borer 2005), and some Romance languages (see next section)). But we have seen that Malagasy does not conform to this nice picture—bare nouns do show general number, but they can take variable scope. Thus for Rullmann and You, general number plays a central role in understanding noun phrase interpretation and structure, but their analysis cannot be extended to account for the Malagasy data.

6.3.3 The view from Romance and Creoles
There is a growing body of literature on bare nouns in Romance languages (e.g. Munn and Schmitt 2005 on Brazilian Portuguese). In a recent article, Espinal (2010) argues that bare nouns in Catalan (22a) and Spanish (22b) are distinct from mass nouns and bare plurals in both interpretation and structure. Consider the following examples (Espinal 2010: 984–985).

(22) a. Tinc pis. [Catalan]
have apartment
‘I am an apartment owner.’ (of one or more apartments)

b. Comprará coche.
buy-fut car
‘(S)he will buy a car.’ (it could be one or more than one)

8 According to Fassi Fehri (2004, 2007), bare nouns in Standard Arabic can take variable scope and certain bare nouns allow for general number interpretations.
She claims that bare nouns lack all functional structure in the DP (they are NPs) and that they denote properties of kinds (not kinds, contra Chierchia). It is this denotation that leads to the number-neutral interpretation. Not surprisingly, bare nouns in Romance take obligatory narrow scope and therefore fit with Rullmann and You’s analysis of bare nouns and general number.

It should be noted that although some Romance languages allow bare noun arguments, their internal and external distribution is highly restricted, thus a direct comparison with Malagasy is not appropriate. And yet, Romance bare nouns otherwise pattern with bare nouns in many other languages in their interpretation (number neutrality and narrow scope). From this perspective, a comparison is warranted.

Turning now to Creoles, a different picture emerges. As carefully documented in the papers in Baptista and Guéron (2007), many Creole languages have relatively unrestricted bare nouns. Crucially for this chapter, most also have general number. Let us consider Cape Verdean Creole, as described by Baptista (2007). This language has definite and indefinite determiners that indicate singular versus plural and there is overt plural morphology on nouns. Bare nouns, however, can receive both definite (anaphoric) and indefinite (novel) interpretations. Moreover, bare nouns show general number effects. Baptista also shows that bare nouns can take both wide and narrow scope (Baptista 2007: 69–70).

[Cape Verdean Creole]

(23) a. Pedru kre kaza k’un/ø bankeru – ma e ka ta da ku el.
   Pedru want marry with a/ø banker but he neg tma get-along with her
   ‘Pedru wants to marry a banker—but he didn’t get along with her.’

   b. Pedru kre kaza k’un/ø bankeru – ma e ka inkontra-l inda.
   Pedru want marry with a/ø banker but he neg meet her yet
   ‘Pedru wants to marry a banker—but he hasn’t met her yet.’

In Haitian Creole, on the other hand, bare nouns can only take narrow scope (Déprez 2004, 2007).

6.4 The emerging typology

In this section, I discuss in more detail the typology of general number that emerges. What I have been looking at are morphosyntactic and semantic correlates of general number. Some of these properties are summarized in Table 6.2 below. I have included languages that have been argued to show general number and considered

---

9 For Espinal, ‘number-neutral’ differs from general number: the former is used for languages that usually express number in the noun phrase. The basic interpretation, however, is the same: the noun is underspecified for number.
the option of marking number on nouns, the presence of determiners (definite and indefinite), the non-plurality effect, and narrow scope (for Persian, see Ghomeshi 2003).

As can be seen from this table, general number does not appear to correlate with any of the morphosyntactic or semantic properties. The closest point of correlation is narrow scope, but as we have already seen, Malagasy and Cape Verdean Creole provide a counterexample.

What can we conclude from the above pattern? Although the result seems negative (general number doesn’t correlate perfectly with anything), I believe this is not surprising. General number is a fairly widespread phenomenon and is manifested in different ways in different languages. Even from the small sample above, we see general number in different language families and otherwise typologically diverse languages. There is also a more general methodological conclusion: while it is tempting to closely tie together morphosyntax and semantics, the mapping between the two is not one-to-one. Form and function are not always perfectly aligned.

<table>
<thead>
<tr>
<th>LANGUAGE</th>
<th># on N&lt;sup&gt;ε&lt;/sup&gt;</th>
<th>DEFINITE DET</th>
<th>INDEFINITE DET</th>
<th>NON-PL EFFECT</th>
<th>NARROW SCOPE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cape Verdean</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Catalan</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Mandarin</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Haitian Creole</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Malagasy</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Malay</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Persian</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

General number and the structure of DP
This chapter pays attention to the syntactic and semantic properties of the plural marker -hâ in Persian, focusing in particular on its definiteness effect on mass nouns. Based on the descriptive generalizations made in this chapter, a Distributed-Morphology based analysis is proposed of number marking in Persian and how it interfaces with definiteness marking. I argue that number morphology in the context of definite noun phrases is syncretic in the sense that it realizes both a definiteness feature \([+\, \text{def}]\) and a number feature \([\pm \, \text{pl}]\).\(^1\)

7.1 Introduction

Plural marking in Persian can encode both definiteness and cardinality. Two particular observations are of interest here. First, in episodic contexts, typical mass nouns in Persian like barq ‘electricity’ and roqan ‘oil’ are optionally marked with the plural marker -hâ. This however does not result in a coerced (count) reading, contrary to the standard assumption (Ghomeshi 2003, Borer 2005, Mathieu 2007). The plural marker -hâ in these examples seems to play the role of a definiteness marker, the same role as the definite article the in the English translations.

(1) barq(-ā) qat’e.
   electricity(-hâ) cut-clt.is.3sg
   ‘The power is shut off.’

(2) roqan(-ā)-ro rixt-am tu qâblame.
   oil(-hâ)-om poured-1sg in pot
   ‘I poured the oil into the pot.’

\(^1\) This phenomenon is well-known in Haitian Creoles (cf. Aboh and Smith, 2009: Part III), and Japanese. Thanks to one of the anonymous reviewers for bringing this to my attention.
Second, in indefinite noun phrases, as in (3), plural marking cannot appear with numerals + classifiers. It is however possible for the plural marker to optionally occur with numerals + classifiers in definite noun phrases such as (4).^2

(3) se-tâ ketâb(-â) xarid-am.
three-cl book(-hâ) bought-1sg
‘I bought three books.’

(4) se-tâ ketâb(-â) ru-ye miz-e.
three-cl books(-hâ) on-ez table-is
‘The three books are on the table.’

Three questions arise from this observation that I will attempt to address in this study: (i) What prevents the plural marker from appearing in indefinite noun phrases involving numerals + classifiers? (ii) What triggers the co-occurrence possibility of plural marking and numerals + classifiers in definite noun phrases? (iii) Why is the plural marker optional in the context of definite noun phrases with numerals + classifiers? It will be shown that the possibility of having the plural marker -hâ on mass nouns in episodic contexts, and on definite noun phrases involving numerals + classifiers is motivated by the same factor: definiteness marking.

Taking into consideration the facts concerning the behavior of number markers in Persian noun phrases, I propose that (i) number and definiteness are syncretic in Persian; and that (ii) number is a modifying feature, in Wiltschko’s (2008) term, not a head of functional projection.

The claim in this chapter is that plural marking on mass nouns and nouns with numerals induces definite readings because in Persian definiteness is bundled with number. The important implication of this claim is that the effect of number marking in Persian is broader compared with English.

7.2 Background on Persian noun phrases

Persian allows nouns to occur bare in argument positions. Bare nouns are construed as either definite (5) or kind-referring (7).^3 Indefinite noun phrases are marked with the indefinite enclitic -i (6), which yields quantitative reading of a noun (Ghomeshi 2003).^4 Define bare nouns in direct object positions, (5b) are distinct from indefinite

---

2 In the context of (3), where the noun phrase se-tâ ketâb 'three books' is in the direct object position, plural marking can appear with numerals + classifiers only if the noun phrase is followed by the direct object marker -râ. Assuming that -râ marks specificity (Karimi 1990) or presupposition/definiteness (Ghomeshi 1996), the optional presence of the plural marker can be due to the definite meaning of the noun phrase.

3 In this chapter, I use kind-referring interchangeably with generic.

4 Ghomeshi (2003: 60) argues that ‘the indefinite marker in Persian is not a marker of cardinality like the numeral one but rather a quantitative indefinite like some, any, and no.’ See the original reference (section 4) for a detailed discussion of the behavior of the indefinite marker -i in Persian and how it differs from that of its counterpart a in English.
(6b) and kind-referring (7b) ones in that they must be followed by the case marker -râ (-ro after vowels and -o after consonants in colloquial Persian).

(5) a. ketâb ru-ye miz-e. b. Sinâ ketâb-o xund.
   book on-3sg table-is Sina book-om read.pst.3sg
   ‘The book is on the table.’ ‘Sina read the book.’

(6) a. ketâb-i gom=-šod. b. ketâb-i xarid-am.
   book-indef lost=became book-indef bought-1sg
   ‘A book was lost.’ ‘I bought a book.’

(7) a. ketâb dust-e mâ-st. b. Sinâ ketâb xund.
   book friend-ez we-is Sina book read.pst.3sg
   ‘Books are our friends.’ ‘Sina read books.’

Ghomeshi (2003, 2008) suggests that the difference in the interpretation of bare nouns parallels a difference in their syntactic category: definite bare nouns are DPs containing an empty D0-head; indefinite bare nouns are QPs; and kind-referring nouns are NPs.

(8) a. DP b. QP c. NP
   NP D NP Q ketâb
   ketâb Ø def ketâb i

7.3 Data

7.3.1 Plural marking on count nouns

Count nouns that instantiate any of the syntactic categories in (8) can be marked with the plural suffix -hâ (-â after consonant-final stems in spoken Persian).5

(9) Definite
   book-pldef on-3sg table-is Sina book-pldef-om read.pst.3sg
   ‘The books are on the table.’ ‘Sina read the books.’

(10) Indefinite
    mâh-e gozašte ketâb-â-ye xub-i xund-am.
    month-past book-pl-3sg good-indef read-1sg
    ‘Last month, I read some good books.’

5 I gloss the plural marker -hâ as pldef when it yields a definite reading and pl when it yields a cardinal reading.
Kind-referring
ketâb-â dust-â-ye xub o qâbel-e + e’temâd-i hast-and.
book-PL friend-PL-EZ good and able-EZ + reliance-INDEF is-3PL.
‘Books are good and reliable friends.’

(source: http://sardargomestan.persianblog.ir/1386/3/)\(^6\)

The interpretation of the kind singular *ketâb* ‘book’ in (7) is different from its plural counterpart in (11): the former is a reference to the kind book as a concept, but the latter refers to all instantiations of the kind book. Combining the plural marker -hâ with the kind noun yields a taxonomic reading, i.e. the plural noun is interpreted as referring to ‘more than one kind’. Thus, the kind plural *ketâb-â* ‘books’ in (11) refers to all sub-kinds of books. I propose that plural marking on kind-referring count nouns introduces a universal quantification over instances of the kind. This proposal is in line with Dayal’s (2004) and Chierchia’s (1998b) arguments that definite generic NPs involve reference to groups.

Assuming that kind-referring nouns in direct object position are construed as forming a complex predicate with the verb (Ghomeshi 2008: 95), plural marking on such nouns is interpreted as referring to the cardinality of the event. In other words, such plural nouns yield adverbial readings.

(12) Sinâ ketâb-hâ xund.
    Sina book-PL read.PST.3SG
    ‘Sina read books a lot.’

(13) ... u bâ gerye harf-hâ-yaš-râ zad o
    s/he with weeping word-PLDEF-CLT.3SG-OM hit.PST.3SG and
    ašk-hâ rixt . . .
    tear-PL shed.3SG
    ‘. . . while weeping, she said her words and shed many tears (= wept a lot) . . .’

Plural count nouns do not occur in indefinite noun phrases introduced by numerals + classifiers. Following Ghomeshi (2003), I assume that this type of noun phrase is generated as a CardP.

(14) a. se-tâ ketâb(*-â) ru-ye miz-e.
    three-CL book(*-PL) on-EZ table-CLT.is.3SG
    ‘Three books are on the table.’

\(^6\) Website links following examples indicate the sources of the data in the examples. These links were functional at the time of press.
However, plural marking can optionally co-occur with numerals + classifiers when the noun phrase is definite. Plural marking in this context functions as a definiteness marker; its presence induces a totality reading for the noun phrase which contrasts with an indefinite cardinality reading. Specifically, the plural marker -hâ in (15) triggers a reading which involves reference to the three books as a whole, not as a set of three individual items. In this sense, -hâ induces a universal quantifier reading, i.e. all the three books. This is empirically supported by the fact that the quantifier har ‘every’ can occur in this noun phrase for emphasis, yielding the same interpretation (cf. (16)).

(15) se-tâ ketâb(-â) ru-ye miz-e.
   three-cl book(-PLDEF) on-ez table-CLT.is.3SG
   ‘The three books are on the table.’

(16) har se-tâ ketâb(-â) ru-ye miz-e.
    every three-cl book(-PLDEF) on-ez table-CLT.is.3SG
    ‘All the three books are on the table.’

In the absence of plural marking, noun phrases involving numerals + classifiers are distinguished by prosody. To elaborate, definiteness in Persian has a prosodic manifestation, affecting prosodic phrasing and prominence. Assuming that a noun phrase with a numeral + classifier is a phonological phrase (φ), an indefinite noun phrase appears as consisting of two phonological words (ω), each with a prominent syllable. In contrast, definite noun phrases with numerals + classifiers involve a phonological phrase which consists of one phonological word, with the prominent syllable in the rightmost element. I assume that this phonological phrasing pattern is a prosodic correlate of definiteness in noun phrases involving numerals + classifiers.

(17) se-tâ ketâb (i) *Indefinite*  (ii) *Definite*
    three-cl book  [[se-tâ]_ω[ketāb]_ω]_φ  [[se-tâ ketāb]_ω]_φ
    ‘three books/the three books’

7.3.2 Plural marking on mass nouns

Plural marking on mass nouns in languages like English and Persian is assumed to be allowed only when the interpretation of the mass noun is subject to coercion—i.e. the count reading of mass nouns (Ghomeshi 2003). As such, plural marking serves the
purpose of portioning out the undivided stuff into conventionalized units or types (Borer 2005, Mathieu 2007). Plural marking on definite mass nouns yields a quantity reading (18) and on kind-referring mass nouns it yields a taxonomic reading (19).

(18) Quantity reading
   a. čây-â-ro gozâšt-am tu sini.
      tea-PLDEF-OM put.PST-1SG in tray
      'I put the teas in the tray.' [i.e. cups/glasses]
   b. in + qad dast=na-kon tu pârč o yax-â-ro bar=na-dâr
      this + much hand=NEG-do.IMP in jar and ice-PLDEF-OM on=NEG-have.IMP
      o bo-xor
      and subj-eat
      ‘Don’t put your hand inside the jar this much and don’t pick the ice(*s) and eat’ [i.e. ice cubes/pieces]

   (source: http://setayesh-koochooloo.blogfa.com/8606.aspx)

(19) Taxonomic reading
   a. čây-â-ye xâreji rang o ta’m-e behtar-i dâr-an.
      tea-PL-ez foreign color and flavor-ez better-INDEF have-3PL
      ‘Foreign teas have better color and flavor.’ [i.e. different kinds/brands]
   b. berenj-â-ye šomâl zud mi-paz-e.
      rice-PL-ez north early DUR-cook-3SG
      ‘Rice(*s) from the north cooks fast.’ [i.e. different varieties]


Pluralizing mass nouns in Persian, however, does not always result in coercion effects. Mass nouns like electricity, water, snow, ice, sand, fog are commonly used with the plural marker -hâ in their typical mass uses, although, crucially, they can also appear as singular (see Tsoulas 2006 for similar cases in Greek). I call mass nouns of this type ‘non-coerced mass nouns’.

(20) barq-â ke mi-r-e, aksolamal-e šomâ či-ye?
   electricity-PLDEF that DUR-go-3SG reaction-ez you what-is.3SG
   ‘What is your reaction when the power shuts off?’

(21) âb-â qat’-e.
   water-PLDEF cut-is.3SG
   ‘The water is shut off.’

(22) barf-â âb=šod. xoršid dune + dune barf-â-ro košt.
   snow-PL water=became.3SG sun seed + seed snow-PLDEF-OM killed.3SG
   ‘The snow melted. The sun killed the snow(*s) [i.e. snowflakes] one by one.’
   (source: http://blog.360.yahoo.com/blog-8bPNlgohfqE6L4IXmYtEHzt?p=383)
As can be seen in the above examples, all the plural non-coerced mass nouns are most felicitously translated into English as definite noun phrases. This indeed reflects the way Persian speakers interpret these plural mass nouns. Each of the examples above presents an episodic context in which the plural mass noun yields an inclusive reading. That is, the reference of the noun is to the totality of the mass that satisfies the description. Thus, in (20) and (21), barq-à ‘the power-pl’ and áb-à ‘the water-pl’ refer to the whole power and water supply of the locations the speakers are talking about. In (22) and (23), barf-à ‘the snow-pl’ and yax-à ‘the ice-pl’ refer to the totality of the snow and the ice that have covered the areas under discussion. mâse-hà ‘the sand-pl’ in (24) is meant to refer to the whole area of the sand on which something is written. Finally, meh-à ‘the fog-pl’ in (25) refers to the whole foggy area that the speaker has passed through.

So far we have seen that plural marking on mass nouns, in addition to its portioning-out function, yields definite descriptions involving inclusiveness. This supports Lyons’s (1999) proposal that definiteness for plural and mass nouns involves inclusiveness (a term due to Hawkins (1978)). I suggest that what the definite article does in the English translations of the non-coerced mass nouns in the above sentences, plural marking does in Persian.

In contexts where non-coerced plural mass nouns are used anaphorically, the previous mention always appears as singular. The following naturally-occurring examples reinforce our claim that plural marking on such mass nouns induces definiteness.

(26) tu in kambud-e áb čerá áb-à-ro hadar=mi-kon-i?
in this shortage-ez water why water-pl-def-om waste=dur-do-2sg
‘With this water shortage, why do you waste the water-pl?’

The definiteness effect of plural marking on mass nouns is further supported by the fact that plural mass nouns cannot be used as predicate nominals (29) and in the existential construction (30).

that-pl water-pl def snow-pl def ice-pl def sand-pl def fog-pl def is-3pl

‘*They are the water/the snow/the ice/the sand/the fog.’

(30) tu satl âb(*-â) bud.
in bucket water(*-pl def) was,3sg

‘There was (*the) water in the bucket.’

Recall that plural marking on kind-referring count nouns yields a universal reading (cf. (11)) to the effect that the plural kind is thought of as referring to all instantiations of the kind. However, kind-referring mass nouns cannot be pluralized, as shown in (31) and (32). I take this as direct evidence that the non-coerced mass nouns in examples (20) to (25) are true mass nouns and that plural marking on kind-referring mass nouns cannot function as what Borer (2005) terms stuff divider. Nor can the plural marker -hâ function as a definite marker because mass nouns, according to Krifka (2004: 22), are generally used as names for atomic kinds, and as such do not need definite markers.

(31) âb(*-â) vase mâ hayáti-ye.
water(*-pl) for we vital-clt.is,3sg

‘Water is vital for us.’

(source: http://saharam.blogsky.com/comments.bs?postid=237)
One interesting observation is that plural marking on non-coerced mass nouns yields a coerced or count reading when it occurs in indefinite noun phrases, as in (33) and (34). From this it follows that definiteness is an added value to the portioning-out function of plural marking in definite noun phrases. The plural marker -hâ in the context of indefinite noun phrases, however, only functions as a number marker.

(33) zemestun-â barf-â-ye sangin-i mi-bârid.
winter-pl snow-pl-ez heavy-indef dur-fell.3sg
'In winters, it snowed heavily.' (lit. some heavy snow-pl fell.) [snow-pl= snowfalls]
(source: http://www.tiknews.org/display/?ID=70669&page=30)

(34) barf-â-ye qašang-i ru deraxt-â jam’-šod-e.
snow-pl-ez beautiful-indef on tree-pl collection=became-pp
'Beautiful snow-pl have stacked on the tree.' [snow-pl= patches of snow]
(source: http://www.flickr.com/photos/7752276@N02/2232431627/)

Relying on our observations, the distribution of the plural marker in Persian can be stated as follows: plural marking can occur in all kinds of count noun phrases and coerced mass noun phrases. However, in non-coerced mass noun phrases and noun phrases involving numerals + classifiers, plural marking is possible only when the noun phrase is definite. Table 7.1 summarizes these generalizations.

<table>
<thead>
<tr>
<th>Table 7.1. Plural marking on Persian noun phrases</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kind</td>
</tr>
<tr>
<td>Count</td>
</tr>
<tr>
<td>Mass</td>
</tr>
<tr>
<td>Coerced</td>
</tr>
<tr>
<td>Non-Coerced</td>
</tr>
<tr>
<td>Numerals + Classifiers</td>
</tr>
</tbody>
</table>

7.3.3 Singular definite marker -é in colloquial Persian

In colloquial Persian, singular definite (sg_{def}) nouns can optionally be marked with the stressed vowel -é (-å before consonants).
As shown in (36) and (37), the optional appearance of this suffix on definite bare mass nouns results in a count reading because the referent is interpreted as being of an understood quantity.

(36) [Watching a snowball melting away]

negā=kon! barf(-é) dâr-e åb=mi-š-e.

Look=do.IMP snow(-sgdef) have-3sg water=dur-become.pres-3sg

‘Look! The snow is melting away.’

(37) [Addressing a boy who is holding a bottle of syrup in his hand]

movâzeb=baš šarbat(-á)-ro na-riz-i ru mobl!

careful=be.IMP syrup(-sgdef)-om neg-pour.pres-2sg on furniture

‘Be careful not to pour the syrup on the furniture!’

We have seen in this section that although morphological marking of mass nouns for number contributes to their definite interpretation, only plural marking on mass nouns can yield a non-coerced reading.

7.4 Definiteness/Number syncretism

So far it has been shown that Persian uses number markers to encode definiteness. This suggests that the semantics of definiteness is bundled together with the semantics of number. Given this, I propose that -hâ functions as a definite/number marker when its number feature is bundled with the definite feature, and as a plural marker when it is specified only with the number feature. Thus, in definite noun phrases, -hâ is a portmanteau morpheme with a feature bundle [+def, +pl]. But in other contexts, -hâ is a homophonous morpheme which represents plural number. This distinction is schematized in (38):

(38) -hâ: plural and definite plural marker

\[
\begin{array}{c}
\text{PL} \\
>1
\end{array} \quad \begin{array}{c}
\text{PL}_{\text{DEF}} \\
>1 \quad \text{DEF}
\end{array}
\]
These two plural morphemes enter into different systems of contrast: \( \text{pl} \) contrasts with the unmarked noun; \( \text{pldef} \) contrasts with definite singular noun (\( \text{sgdef} \)) realized as either -\( \ddot{e} \) or a null morpheme.

Based on this number system, definite noun phrases are distinguished from other types of noun phrases by the feature composition of their number markers. If the feature [\text{def}] is bundled with a value of the feature [\text{pl}], the number marker triggers a definite reading. If not, the noun phrase is ambiguous between being indefinite and kind-referring.

### 7.5 Categorial identity of Persian plural marker

Having established that definiteness marking in Persian can be implemented via the plural marker, we now need to determine the categorial nature of the plural marker and how it is accommodated in the Persian syntax.

**7.5.1 Persian plural marking does not instantiate num**

Ghomeshi (2003) argues that plural marking in Persian, contra English, is licensed within D/QPs rather than NumPs. She assumes that DPs involve definite NPs and QPs involve indefinite NPs. Thus, Persian is different from English in that it lacks a NumP projection.

(39) a. English  
   b. Persian

\[
\begin{align*}
\text{D/QP} & \quad \text{D/QP} \\
\text{...CardP...} & \quad \text{...CardP...} \\
\text{...NumP...} & \quad \text{...NP...} \\
\text{...NP...} & \quad \text{...NP...}
\end{align*}
\]

(Ghomeshi 2003: 71.70)

By assuming that Persian lacks NumP projection, Ghomeshi is implicitly suggesting that in Persian, number in general and the plural marker in particular do not function as an inflectional category. In the next section, I will address the categorial status of the plural marker in Persian.

---

6 See Wiltschko (2008) for a distinction between inflectional vs. non-inflectional number marking.
Building on distributional properties of plural markers in English and Halkomelem, Wiltschko (2008) proposes that plural markers across languages can be either a projecting functional head (as in English) or a modifier adjoined to a head (as in Halkomelem). She proceeds to suggest that plural markers that merge as heads have the syntax of heads and modifying plural markers will have the syntax of adjuncts. She puts forward a number of criteria as diagnostics for distinguishing the categorial status of plural markers as heads vs. modifiers. According to Wiltschko (2008), modificational plural markers ‘[…] do not trigger agreement, their absence is not associated with a specific meaning but instead is truly unmarked, […] and do not allow for form-meaning mismatches’ (p. 688). Let us now apply these three diagnostics to the Persian plural marker.8

First, plural marking in Persian does not trigger agreement. This is shown in (40) where demonstratives do not agree with the head noun in number.9

\[
\begin{align*}
\text{(40)} & \quad \text{in(\text{-}â)} \quad \text{pesar-â} \quad \text{mi-dov-and.} \\
& \quad \text{this(\text{-}PL_{(DEF)})} \quad \text{boy-PL_{DEF}} \quad \text{dur-run.PRES.3PL} \\
& \quad \text{‘These boys are running.’}
\end{align*}
\]

Second, in Persian, unlike English, the unmarked (singular) noun, e.g. ketāb ‘book/books’, is compatible with a plural interpretation; therefore, Persian has what Corbett (2000) calls general number.

Third, Wiltschko predicts that form-meaning mismatches arise if plural marking is inflectional (i.e. a syntactic head). She attributes the form-meaning mismatches in examples of pluralia tantum in English to the head status of plural marking in this language. Persian does not have any cases of pluralia tantum. Nouns instantiating pluralia tantum in English, e.g. scissors, pants, glasses, etc., always appear singular in Persian. When pluralized, they are construed as referring to non-atomic sets of pairs.

\[
\begin{align*}
\text{(41) a. qeyči-hâ} & \quad \text{scissors-PL} \\
\text{‘pairs of scissors’} \\
\text{b. šalvâr-â} & \quad \text{pants-PL} \\
\text{‘pairs of pants’} \\
\text{c. eynak-â} & \quad \text{glasses-PL} \\
\text{‘pairs of glasses’}
\end{align*}
\]

---

8 Wiltschko (2008) introduces other criteria as diagnostics for modificational plural markers. I do not use those diagnostics because they are too complicated and not fine-grained enough to capture the choice of the plural marker in Persian (see Wiltschko’s work for the details of those diagnostics).

9 Another instance of lack of agreement in Persian is constructions with inanimate plural subjects. In such constructions, the verb does not agree with the subject in number, e.g.

\[
\begin{align*}
\text{(i) barg-â} & \quad \text{rixt.} \\
\text{leaf-PL} & \quad \text{fell.3SG} \\
\text{‘Leaves fell.’}
\end{align*}
\]

I do not consider this as viable evidence in support of the modificational status of the plural marker in Persian because I believe lack of agreement in such contexts is triggered by ‘the degree of volition or control of the verbal action by the subject’ (Sedighi 2005: 188). According to Sedighi, agenthood, not subjecthood, is responsible for verbal agreement in Persian. Therefore, because inanimate plural subjects are not agents, they do not agree with verbs (see the original work for further details).
Based on these observations, I propose that the Persian plural marker is a modifier adjoined to the head noun. I follow Marantz (1997, 2001) in assuming that the head noun, as a lexical category, is syntactically derived by merging a category-neutral root with the category-defining functional head $n$.

(42) **Persian plural marker as a modifier**

$$
\begin{array}{c}
\sqrt{\text{ROOT}} \\
\quad n \\
\quad \# \left[ +\text{pl} \right] \left[ +\text{def} \right] \\
\quad n \\
\end{array}
$$

7.6 Analysis

In this chapter, I assume Distributed Morphology (Halle and Marantz 1993, and subsequent works) as my theoretical framework. Under DM, word structure and phrase structure are derived solely from syntax. Syntactic structures are input to morphology, where phonological exponents are inserted into terminal nodes of the syntactic structures late in the derivation.

7.6.1 The syntax of number/definiteness marking in Persian

Building on the descriptive generalizations of number marking in Persian, I propose that the following syntactic structures are input to morphology. I use the label $nP_{\text{def}}$ for definite noun phrases—i.e. those involving the feature bundle $[\pm \text{pl}, +\text{def}]$.

(43) a. Kind-referring nouns  
    
    $$
    \begin{array}{c}
    nP \\
    \quad n \\
    \quad \# \left[ +\text{pl} \right] \\
    \quad \left< \sqrt{\text{ROOT}} \right> \ldots \\
    \sqrt{\text{ROOT}} \\
    \end{array}
    $$
    
    b. Definite nouns  
    
    $$
    \begin{array}{c}
    nP_{\text{def}} \\
    \quad n \\
    \quad \# \left[ +\text{pl} \right] \left[ +\text{def} \right] \\
    \quad \left< \sqrt{\text{ROOT}} \right> \ldots \\
    \sqrt{\text{ROOT}} \\
    \end{array}
    $$
    
    c. Indefinite nouns  
    
    $$
    \begin{array}{c}
    QP \\
    \quad nP \\
    \quad \# \left[ +\text{pl} \right] \\
    \quad \left< \sqrt{\text{ROOT}} \right> \ldots \\
    \sqrt{\text{ROOT}} \\
    \end{array}
    \begin{array}{c}
    Q \\
    \quad -i \\
    \end{array}
    \begin{array}{c}
    \sqrt{\text{ROOT}} \\
    \quad n \\
    \quad \left< \sqrt{\text{ROOT}} \right> \ldots \\
    \end{array}
    $$
The phonological exponents of morphemes are inserted at Vocabulary Insertion. Assuming that \( pl \) and \( pl_{\text{def}} \) are two independent morphemes with different feature specifications, it is expected that Vocabulary Insertion assigns different Vocabulary Items (VIs) to each. As we discussed earlier, the homophonous \(-hâ\) is the phonological exponent for both of these two morphemes, as schematically shown in (44) and (45).

(44) **VI for the definite plural marker in Persian**

\[-hâ \leftrightarrow [+pl] +\text{def}\]

(45) **VI for the plural marker in Persian**

\[-hâ \leftrightarrow [+pl]\]

(46) **VI for the definite singular marker in Persian**

\[-\emptyset / -\text{ê} \leftrightarrow [-pl] +\text{def}\]

In (44), we are dealing with a portmanteau (syncretic) plural morpheme with the feature bundle \([+\text{def}, +\text{pl}]\) which occurs only in definite noun phrases. In (45), we have the VI for the plural marker which can occur in contexts other than definite noun phrases. The VI for the definite singular is represented in (46). This morpheme can have either a null exponent or be realized as \(-\text{ê}\).

### 7.6.2 Plural marking in noun phrases with numerals + classifiers

Given our analysis, in this section I try to provide answers to the questions raised in the introduction regarding the behavior of the plural marker in noun phrases with numerals + classifiers. I follow Ghomeshi (2003) in assuming that noun phrases with numerals + classifiers are generated as CardPs. I take (47) as the representative structure of indefinite noun phrases with numerals + classifiers:

(47) CardP

```
<table>
<thead>
<tr>
<th>CardP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Card_{[+pl]} CLP</td>
</tr>
<tr>
<td>se 'three' CL nP</td>
</tr>
<tr>
<td>-tâ</td>
</tr>
<tr>
<td>n (\sqrt{\text{KETÂB}}) {\text{book}}</td>
</tr>
<tr>
<td>n #_{[+pl]} \emptyset</td>
</tr>
<tr>
<td>\sqrt{\text{KETÂB}} n</td>
</tr>
</tbody>
</table>
```
As illustrated above, I propose that the head of CardP is inherently specified with the feature [+pl] and that this feature prevents the spell-out of the plural marker, which also bears the feature [+pl] lower in the structure. In other words, in such contexts, the realization of the feature [+pl] inherently specified on numerals renders the spell-out of another [+pl] feature redundant. This instantiates redundancy avoidance and can be stated as the following economy constraint, similarly proposed by Ortmann (2000):\(^{10}\)

\[(48)\] **Principle of Avoiding Redundant Plural Marking**

There is no more than one realization of [+pl] within the noun phrase.

My discussion so far has provided the answer to the first question regarding what prevents the plural marker from appearing in indefinite noun phrases involving numerals + classifiers. The second and third questions are concerned with plural marking in the definite noun phrases with numerals + classifiers. This kind of noun phrase is generated as a CardP that takes a definite nP as its complement.

\[(49)\]

\[
\text{CardP} \\
\text{Card}_{[+pl]} \quad \text{CLP} \\
\text{se} \quad \text{CL} \quad n\text{P}_{\text{def}} \\
\text{‘three’} \quad -\text{tâ} \quad n \\
\sqrt{\text{KETÄB}} \quad (\sqrt{\text{KETÄB}}) \\
\sqrt{\text{KETÄB}} \quad n \\
\]  

The node that is realized as -hâ in (49) is characterized by a feature bundle containing the features [+pl] and [+def]. I suggest that in this context, -hâ spells out [+def], not [+pl]. In other words, the co-occurrence possibility of plural marking and numeral is motivated by a morphological requirement that [+def] be spelled out. This is the answer to the second question.

The third question concerning the optionality of plural marking in the presence of numerals + classifiers is perhaps the most troubling question for which I cannot provide an answer within DM. However, I am going to propose an account inspired by the constraint ranking system of Optimality Theory (Prince and Smolensky 1993) which addresses all the issues surrounding the occurrence of the plural marker in noun phrases with numerals + classifiers in Persian.

\(^{10}\) Ortmann’s (2000) Principle of Economic Plural Marking (PEPL) accounts for the disallowance of several noun-phrase internal realizations of plural in Hungarian and many other languages including Persian. He calls these languages Type Hungarian and compares them with languages of Type English, in which several overt realizations of plurality are possible. To capture the difference between these two language types, he proposes an Optimality-Theoretic analysis in which plural marking in each of these language types occurs under a different constraint ranking. In Type Hungarian languages, Ortmann suggests, PEPL is undominated (see the original work for more details).
I propose that the appearance of the plural marker \(-hâ\) in noun phrases involving numerals + classifiers is determined by economy through the interaction between the following constraints:

(50)  
(i) \([+pl]\) must be phonologically realized.  
(ii) \([+\text{def}]\) must be spelled out.  
(iii) Avoid redundancy in number marking.

In indefinite noun phrases, the presence of the plural marker results in ungrammaticality because economy disallows the co-occurrence of the plural marker and numeral + classifier. Thus, I propose that in deriving indefinite noun phrases with numerals, constraint (iii) outranks constraint (i). In definite noun phrases, the plural marker is optional because constraints (ii) and (iii) are tied. In other words, both possibilities are equally economical.

7.6.3 Plural marking on mass nouns

As I showed in section 7.3.2, plural marking on mass nouns like barq ‘power’, mâse ‘sand’, and barf ‘snow’ results in a non-coerced reading only when it occurs in definite noun phrases. In indefinite noun phrases, however, the occurrence of \(-hâ\) results in a coerced (count) reading (cf. (33) and (34)). The difference in the interpretations of \(-hâ\) can be accounted for if we assume that the node for \(-hâ\) in definite and indefinite noun phrases has different morphosyntactic feature specifications. I propose that the non-coerced reading of mass nouns is obtained when \(-hâ\) is inserted for \(\text{pl}_{\text{def}}\), while the coerced reading is obtained when \(-hâ\) is inserted for \(\text{pl}\). The following syntactic structures represent the contexts in which plural marking occurs with mass nouns:11

(51) a. Non-coerced reading  

b. Coerced reading

\[
\begin{align*}
\sqrt{\text{ROOT}} & \quad \text{nP}_{\text{def}} \\
\sqrt{\text{ROOT}} & \quad \text{n} \\
\sqrt{\text{ROOT}} & \quad \text{\#}_{[+\text{pl}]}^{[+\text{def}]} \quad \text{\langle \sqrt{\text{ROOT}} \rangle \ldots} \\
\sqrt{\text{ROOT}} & \quad \text{n} \\
\sqrt{\text{ROOT}} & \quad \text{n} \\
\sqrt{\text{ROOT}} & \quad \text{\#}_{[+\text{pl}]} \quad \text{\langle \sqrt{\text{ROOT}} \rangle \ldots} \\
\sqrt{\text{ROOT}} & \quad \text{n}
\end{align*}
\]

where \(\sqrt{\text{ROOT}} = \{\sqrt{\text{BARQ}}, \sqrt{\text{MÂSE}}, \sqrt{\text{BARF}}, \text{etc.}\}\)

‘power’ ‘sand’ ‘snow’

11 I have no explanation to offer for the correlation between being non-coerced vs. coerced and being definite vs. indefinite.
Assuming that in (51a) -hå spells out the [+def] feature, the question arises as to how the feature [+pl] contributes to the semantics of non-coerced mass nouns, given that they are construed as unstable entities (i.e. vague in number). With reference to the fact that the plural marker appears on non-coerced mass nouns only in episodic contexts, I hypothesize that, as far as the interpretation of number is concerned, -hå functions as a precisifier. I take -hå as an operator that takes a vague expression (e.g. a true mass noun) and returns a precise one, relative to the context of utterance (see Chierchia 2009 for a semantic model).

We have seen in this chapter that the occurrence of the plural marker -hå on definite count/mass nouns induces both number and definiteness. This indicates that what licenses plural marking in Persian noun phrases is the morphosyntactic feature composition of the node for number markers, not the noun’s count/mass status.

7.7 Conclusion

In this chapter, I have argued for an analysis of definiteness marking via Persian number markers. I have proposed that

- The definite interpretations associated with number markers are derived from their morphosyntactic feature geometry and the syntactic structure where they occur;
- In the Persian number system, number and definiteness are syncretic;
- Number in Persian is a modifying feature in the sense that it does not project as a functional head.

In the light of the analysis presented in this chapter, it is fair to say that the effect of number marking in Persian is broader compared with English because in Persian definiteness is bundled with number.
Aspectual effects of a pluractional suffix: Evidence from Lithuanian*

SOLVEIGA ARMOSKAITE

8.1 The problem: -inė- is not a dedicated aspectual morpheme

This section introduces the key problem, namely the puzzling data of the suffix -inė-traditionally labeled as aspectual. First, I state my assumptions about aspect (8.1.1). Next, I show how, based on these assumptions some data with -inė- is difficult to account for (8.1.2).

8.1.1 Assumptions about aspect

Verbal aspect has to do with the time of an event. There are two types of aspect: ‘outer aspect’ (also known as ‘grammatical aspect’) and ‘inner aspect’ (also commonly labeled as ‘Aktionsart’ or ‘lexical aspect’).

Inner aspect classifies verbal predicates into accomplishments (draw a circle), achievements (win), statives (love), and activities (run) (Vendler 1967). For the purposes of this study, I assume that [± telic] and [± stage] are the two crucial aspecual properties that distinguish between the four classes (Rothstein 2006:12).

Telicity (from Greek telos ‘end’) refers to the endpoint of an event. If the event has an endpoint, then it is telic. Atelic events do not have an endpoint. Achievements and accomplishments are considered telic, while statives and activities are considered atelic events.¹

The property of having stages captures whether the event (e’) expressed by a verbal predicate has subevents (e) or not. A predicate has stages if e is a stage of

* This chapter has benefited from discussions with: two anonymous reviewers; Jurgys Pakerys; Marina Sherkin-Lieber; Martina Wiltschko; University of British Columbia, Department of Linguistics students and faculty; University of California, Department of Linguistics students and faculty; and the participants of the Mass-Count Workshop 2009, Toronto. All remaining errors are mine.

¹ Cross-linguistically, what is considered, for example, an accomplishment in a language like English may not hold for a language like Squamish. See Bar-el (2005) for a detailed discussion.
e’ when e develops into e’. Statives and achievements do not contain stages, while accomplishments and activities have stages.

Based on the two criteria, the four classes are often arranged as in Table 8.1 (based on, e.g., Rothstein 2004, among others):

<table>
<thead>
<tr>
<th>Table 8.1. Aspectual classes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aspectual class</td>
</tr>
<tr>
<td>State</td>
</tr>
<tr>
<td>Activities</td>
</tr>
<tr>
<td>Achievements</td>
</tr>
<tr>
<td>Accomplishments</td>
</tr>
</tbody>
</table>

8.1.2 The problem with -inė-

The Lithuanian suffix -inė- is considered aspectual: it cancels the endpoint of telic events (Ambrazas et al. 1997:236). For example, we know that užbėgo ‘ran up’ is a telic predicate because it is selected by an adverbial ‘in ten minutes’ that picks out telic events (1a):

(1) a. Pelė užbėgo į kalną per dešimt minučių.
   pel-ė už -bėg-o į kaln-ą per dešimt minuč -ių
   mouse-NOM.SG pref-run-PAST on hill-ACC.SG in ten minute-ACC.PL
   ‘The mouse ran up the hill in ten minutes.’

   b. *Pelė užbėgo į kalną kiaurą dieną.
   pel-ė už-bėg-o į kaln-ą kiaurą dien-ą
   mouse-NOM.SG pref-run-PAST on hill-ACC.SG whole-ACC.SG day-ACC.SG
   Intended: ‘The mouse ran up the hill the whole day.’

   In contrast, an adverbial like ‘the whole day’ picks out atelic events, therefore (1b) is ungrammatical. Once -inė- is added, the telic predicate becomes atelic and can be picked out by adverbials sensitive to atelicity (1a versus 2):

(2) Pelė užbėginėjo į kalną kiaurą dieną.
   pel-ė už-bėg -inė -j-o į kaln-ą kiaurą dien-ą
   mouse-NOM.SG pref-run-SUF-PAST on hill-ACC.SG whole day-ACC.SG
   ‘The mouse kept running up the hill all day long.’

2 The online searchable Lithuanian dictionary http://www.lkz.lt is the main source of data. When in doubt, linguistically naïve Lithuanian native speakers have been consulted. Special thanks to Dr. Jurgis Pakerys (Vilnius University) for detailed comments and suggestions on Lithuanian data.
However, there exists an entire subset of data where the role of -\textit{inė}- cannot be accounted from the aspectual point of view. If -\textit{inė}- were a dedicated atelic affix, the following data would be puzzling:

(3) a. Pelė bėgo \textit{kiaurą dieną}.
   \begin{tabular}{l}
   pel-ė  & bėg-o  & kiaurą  & dien-ą \\
   mouse-NOM.SG & run-PAST & whole day- ACC.SG
   \end{tabular}
   'The mouse ran all day long.'

b. Pelė bėginėjo \textit{kiaurą dieną}.
   \begin{tabular}{l}
   pel-ė  & bėg-inė-o  & kiaurą  & dien-ą \\
   mouse-NOM.SG & run -SUF-PAST,3SG & whole day- ACC.SG
   \end{tabular}
   'The mouse scurried around all day long.'

Given that the predicate in (3a) is already atelic, the role of -\textit{inė}- in (3b) would be unclear if it were considered an atelic aspectual marker. This lack of contrast in telicity indicates that marking of aspectuality is not the primary function of -\textit{inė}-. However, if -\textit{inė}- is viewed as a pluractional maker, then the contrast in (3a–b) is clear: a plurality of events is introduced (Bar-el 2008 reaches a similar conclusion for Squamish). Therefore I argue that -\textit{inė}- is not an aspectual suffix, and its aspectual effects are only epiphenomenal.

An anonymous reviewer has pointed out that there is a tradition to treat verbal plurality as a subcase of lexical aspect (e.g. Cusic 1981, Dressler 1968, among others). On the one hand, data like in (3) pose a problem to the traditional view that plural event interpretation is only a subcase of aspect. On the other hand, there are alternative views on the aspect-event number relation. For example, Ferreira (2005) draws on the parallelism between DPs and VPs with respect to number (see also Bar-el 2008). He posits that VP-denotations may contain atomic as well as non-atomic events. Number morphemes combine with bare VPs (Ferreira 2005:97), and can be of the following logical form:

(4) a. $[\text{TP Past [AspP Imp [vp-sg sg [vp John paint the house]]]]}$

b. $[\text{TP Past [AspP Imp [vp-pl pl [vp John paint the house]]]]}$

### 8.2 Suffix -\textit{inė}- is Number

In this section I argue that the Lithuanian suffix -\textit{inė}- marks plural events in the verbal domain. I first discuss evidence for its status as Number (8.2.1). Then I show that -\textit{inė}- is a modifier (8.2.2). Lastly, I show that it modifies verbal roots (8.2.3).

\footnote{For more on the aspect-pluractionality relation, see concerns raised in Laca (2006), footnote 10.}
8.2.1 -inė-: a plural suffix of events

I show that the primary function of -inė- is to mark plurality of events (8.2.1.1). I also discuss how its distributive interpretation comes about. Distributivity arises as result of interaction between the lexical meaning of the root and argument structure (8.2.1.2).

8.2.1.1 Evidence for -inė- as a plural marker I propose that -inė- is a plural marker of events (see Yu 2003 for similar proposal for Chechen, Bar-el 2008 on Squamish). The contrast in interpretation between the simplex verb and a verb affixed with -inė- is the first indication of plurality of events. The simplex verb cannot be interpreted as denoting frequent or multiple events (5a), while the affixed verb cannot be interpreted as denoting a single event (5b).

   Jonas važ-ia-v-o į Toront -ą
   John travel-SUF-3SG.PAST to Toronto-ACC.SG
   ‘John was traveling/travelled to Toronto.’
   ≠ ‘John made frequent/repeated trips to Toronto.’

b. Jonas važinėjo į Torontą.
   Jonas važ -inė-jo į Toront-ą
   Jonas travel-SUF-3SG.PAST to Toronto-ACC.SG
   ‘John made frequent/repeated trips to Toronto.’
   ≠ ‘John was traveling/travelled to Toronto.’

Selectional restrictions with adverbials support the contrast in meaning. Adverbials that pick out single atomic events are infelicitous with -inė-. For example, adverbial tris kartus ‘three times’ is felicitous with a simplex verb (6a), while it is semantically unacceptable with an -inė- suffixed verb (6b):

(6) a. Jonas važiavo į Torontą tris kartus.
   Jonas važ -ia-v-o į Toront-ą tris kart-us
   John travel-SUF-3SG.PAST to Toronto-ACC.SG three times- ACC.PL
   ‘John was traveling/travelled to Toronto three times.’

b. #Jonas važinėjo į Torontą tris kartus.
   Jonas važ-inė-jo į Toront-ą tris kart-us
   Jonas travel-SUF-3SG.PAST to Toronto-ACC.SG three times- ACC.PL
   Intended: ‘John made frequent/repeated trips to Toronto three times.’

Similarly, a simplex verb is felicitous with the adverbial vieną kartą ‘once’ (7a); while -inė- is infelicitous with it (7b):

4 For an alternative view, see van Geenhoven (2004), who argues for atelicity being the source of pluractionality.
(7)  a. Vieną kartą Jonas važiavo į Torontą.
    vieną kartą Jonas važiavo į Torontą
    one time ACC.SG John travel SUF-PAST to Toronto ACC.SG
    ‘Once, John was traveling to Toronto.’

    b. # Vieną kartą Jonas važinėjo į Torontą.
    vieną kartą Jonas važinėjo į Torontą
    one time ACC.SG John travel SUF-PAST to Toronto ACC.SG
    Intended: ‘Once John made frequent/repeated trips to Toronto.’

However, both the simplex verb and the verb affixed with -inė- are felicitous with adverbials that pick out multiple undefined number of events. Dažnai ‘often’ is one example of such an adverbial:

(8)  a. Jonas dažnai važiavo į Torontą.
    Jonas dažnai važiavo į Torontą
    John often travel SUF-PAST to Toronto ACC.SG
    ‘John was traveling/travelled to Toronto often.’

    b. Jonas dažnai važinėjo į Torontą.
    Jonas dažnai važinėjo į Torontą
    John often travel SUF-PAST to Toronto ACC.SG
    ‘John often made frequent/repeated trips to Toronto.’

Even with ‘often’, the iterativity contrast between the two verbal forms is retained. In (8a), trips occur often; in (8b) the repeated trips occur often.

8.2.1.2 Environments where -inė- attains distributive interpretation Verbs are considered distributive when an event sequentially affects one object (or subject) after another. The verbs may be interpreted as distributive as a result of being affixed with -inė-. However, while the distributive interpretation is possible, it is not necessarily obligatory, because many verbs with -inė- do not require an object. For example, gulti ‘lay down’ is such a verb:

(9)  Ligonė iš silpnumo gulėjo.
    Ligonė iš silpnumo gulėjo
    patient NOM.SG from weakness GEN.SG lay SUF-PRES3SG
    ‘The patient keeps laying down due to weakness.’

This example does not have a distributive interpretation. The subject is in singular and the verb does not have an object to affect. With a slight change of context and argument structure, the distributive interpretation is made available:
Kiškiai gulinėja po javus.
hare-NOM.PL lay-SUF-PRES.3SG over wheat-ACC.PL

‘The hares keep laying down all over wheat field.’

The subject and the object are in the plural, and the verb with -inė- acquires distributive meaning of multiple events of rabbits flattening wheat in multiple places all over the wheat field.

There are verbs that cannot be used with -inė- other than with distributive interpretation. The verb mirti ‘die’ is an example of such a verb.

a. Žmonės labai mirinėja.
human-NOM.PL very die-SUF-PRES.3PL

‘People are dying out one by one.’

≠ ‘People die repeatedly.’

Context: bubonic plague

b. Žmogus labai mirinėja.
people-NOM.SG very die-SUF-PRES.3SG

Intended: ‘A man dies repeatedly.’

In (11a), the primary interpretation is that of distributivity, where scores of people are taken by plague one by one. The intended interpretation in (11b) is semantically strange due to the subject being in singular which excludes a distributive reading.

In sum, I have shown that although -inė- contributes to distributive interpretation, it is neither the primary nor uniformly available property of the suffix. Moreover, it is sensitive to the lexical meaning of the verb.

8.2.2 Evidence for the modifier status of -inė-

The proposal that -inė- is a modifier builds on the following assumptions. First, I assume that heads are obligatory, while modifiers are optional. Second, modifiers can be paraphrased, while heads cannot be paraphrased. And lastly, modifiers are sensitive to the lexical meaning of the linguistic object being modified.

Number in the nominal domain is obligatory, i.e. it is a head. It is always obligatorily marked on every nominal, and it cannot be paraphrased:

Based on example (10), a reviewer observes that the preposition po may be the sole source of distributivity. If po were the sole source of distributivity, we would also expect po in (11), yet it is not required. Therefore one can argue that -inė- may also be a source of distributive interpretation.
(12) a. dantis  
   dant-is  
   tooth-MASC.NOM.SG  
   ‘tooth’
   dantys
   dant-yis  
   tooth-MASC.NOM.PL  
   ‘teeth’

b. *dant
   tooth
   Intended: tooth

c. *daug dant
   many tooth
   Intended: many teeth

In (12a), dant ‘tooth’ is given in the plural and singular form, while an unmarked form is ungrammatical (12b). An attempt to add a quantifier daug ‘many’ results in ungrammaticality (12c). In contrast, Number in the verbal domain is a modifier and it is optional (see Bar-el 2008 for similar facts in Squamish), and can be paraphrased:

(13) a. sukti
   suk-ti
   turn-INFL
   ‘to turn’
   b. sukinėti
   suk-inė-ti
   turn-SUF-INFL
   ‘to turn many times/directions’

c. sukti daug kartų į visas puses
   suk-ti  daug kart-ų į vis-as  pus-es
   turn-INF  many time-GEN.PL  in all-ACC.PL  direction-ACC.PL
   ‘to turn many times/directions’

As (13a–b) show, whether -inė- is present or not, the grammaticality is not affected. And, unlike with nouns, a periphrastic construction with the quantifier daug ‘many’ is possible (13c).

The contrast between Number heads and Number modifiers is also upheld with respect to the inherent meaning of the linguistic object. Number on nominals is an obligatory lexical meaning, as is expected when Number is a head. There simply are no lexical entries that lack Number marking. Verbal Number is sensitive to the inherent meaning of verbal roots, i.e. it behaves like a modifier. For example, -inė- is felicitous with verbs of motion (14), or with verbs of separating and disassembling (14), and so on (verb classes based on Levin 1993).

(14) -inė- as a modifier of a motion verb

a. važiuoti
   važ-iuo-ti
   travel-SUF-INF
   ‘travel’

b. važinėti
   važ -inė-ti
   travel-SUF-INF
   ‘travel repeatedly’

6 The φ features—gender, number and case—are fused in Lithuanian. I assume that the three are obligatory. Note that gender is not glossed in other examples in the interest of space, given that it is irrelevant to the discussion at hand.
(15) -inė- as a modifier of a separating and disassembling verb
   a. kalti     b. kalinėti
      kal-ti    kal-inė-ti
      hammer-INF hammer-SUF-INF
      'hammer'   'hammer repeatedly, lightly'

However, there are verbs that -inė- does not attach to. For example, verbs of social interaction (16) and dining (17) do not take on the suffix.

(16) -inė- as a modifier of a social interaction verb
   a. ginčyti     b. *ginčinėti
      ginč -y -ti    ginč-inė-ti
      argue-SUF-INF argue-SUF-INF
      'argue'        Intended: 'argue repeatedly'

(17) -inė- as a modifier of dining verbs
   a. pusryčiauti b. *pusryčinėti
      pusryč -iau-ti pusryč-inė-ti
      breakfast-SUF-INF breakfast-SUF-INF
      'have breakfast' Intended 'have breakfast repeatedly'

Examples as in (16b–17b) are plausible: there is nothing inherently wrong with numerous arguing or breakfasting events. Yet an attempt to derive these meanings with the suffix -inė- is ungrammatical. I take this as yet another indication that the use of -inė- is idiosyncratic, i.e. that of a modifier rather than a head.

Thus, I have shown that based on the three criteria above, the behavior of -inė- is consistent with a modifier, and not with a head. The contrast is summed up in Table 8.2 below.

<table>
<thead>
<tr>
<th>Number</th>
<th>Obligatory</th>
<th>Paraphrase</th>
<th>Semantic restrictions</th>
</tr>
</thead>
<tbody>
<tr>
<td>head</td>
<td>✓</td>
<td>×</td>
<td>✓</td>
</tr>
<tr>
<td>modifier</td>
<td>×</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

8.2.3 Evidence for the root level merge of -inė-

In the previous section I have shown that -inė- is a modifier. I will now show the evidence for -inė- attaching at the level of verbal roots. For the purposes of the discussion, I assume that roots are of two types: inherently categorized or category-
neutral (for discussion of root types see Armoskaite 2011 and references therein). Crucially, -inė- modifies verbal roots but does not derive verbs.

Consider a category-neutral root švies- ‘light’ in its distinct incarnations of adjective (18a), noun (18b), and verb (18c).

(18) a. šviesus švies
švies-us švies-i
light-NOM.SG.MASC light-NOM.SG.FEM
‘light’ ‘light’

b. šviesa
švies-a
light-NOM.SG.FEM
‘the/a light’

c. šviesti
švies-ti
light-INF
‘to emanate light’ \textit{intransitive}
‘to shed light’ \textit{transitive}

The morphology in these examples is inflectional, not derivational. In other words, these affixes cannot be attached to just any root to derive the distinct lexical categories. For example, it is not possible to take an inherently categorized root like bais- ‘terrible’, attach an infinitive suffix and derive a verb (19a); the derivation is only possible when a causative suffix -in- is introduced (19b).

(19) a. *baisinėti
bais-inė-ti
terrible-SUF-INF
Intended: ‘to scare many times’

b. baisinti
bais-in-ti
terrible-CAUS-INF
‘to cause to be terrified’

Based on examples like in (18–19) I assume that category neutral roots may be of any category without any overt derivational morphology, while this is not the case with the inherently categorized roots.\footnote{Categorization of Lithuanian roots requires more study. See Armoskaite (2011) for a discussion on different types of Lithuanian roots from a generative perspective.} Thus, in (20), when -inė- attaches to švies- ‘light’,
it modifies a verbal root; an attempt to keep -\textit{inė}- with the nominal and adjectival incarnations of the same root is ungrammatical (20b–c).

(20) a. :\ Šviesinētī  \hspace{1cm} \sqrt{v} + v + \textit{-inė-}  
   Švies-\textit{inė}-ti  
   light-SUF-\textit{INF}  
   ‘to emanate light (in bursts)’ \textit{intransitive}\n   ‘to shed light (in bursts)’ \textit{transitive}\n
   b.  *Šviesinēa  
   Švies-\textit{inē}-a  
   light-SUF-NOM.SG.FEM

   c.  *Šviesinēus  
   Švies-\textit{inē}-us  
   light-SUF-NOM.SG.MASC

Close proximity to the root is a requirement. Verbal roots take other affixes, such as e.g. the infinitive suffix -\textit{ti} or reflexive suffix -\textit{s(i)}:

(21) a.  suk\textit{tis}  
   suk-\textit{ti} -\textit{s}  
   turn-\textit{INF-REFL}  
   ‘turn oneself’

   b.  suk\textit{inētis}  
   suk-\textit{inē} -\textit{ti} -\textit{s}  
   turn-SUF-\textit{INF-REFL}  
   ‘turn oneself many times/directions’

However, when -\textit{inē}- is introduced, it attaches closest to the root, and it is ungrammatical to attach it elsewhere:

(22) a.  *Sukti\textit{inēs}  b.  *Sukti\textit{sinē}  
   suk-\textit{ti} -\textit{inē} -\textit{s}  \hspace{1cm} suk-\textit{ti} -\textit{s} -\textit{inē}  
   turn-\textit{INF-SUF-REFL} \hspace{1cm} turn-\textit{INF-REFL-SUF}

Another argument for root level suffixation is the idiosyncrasy of meanings that -\textit{inē}- allows:

(23) a.  tū\textit{pti}  
   tūp-\textit{ti}  
   squat-\textit{INF}  
   ‘to squat’
b. tupinėti apie svečius
   tup-inė-ti apie sveč-ius
   squat-SUF-INF around guest-ACC.PL
   Lit.: ‘to squat around guests’
       ‘to fuss about guests’

c. dantis
dant-is
tooth-NOM.SG
‘a tooth’
d. dantinėti
dant-inė-ti
tooth-SUF-INF
Lit.: ‘take apart by teeth’
       ‘to gossip about’

We can see that the presence of -inė- yields a figurative interpretation of squatting (23b); while in (23d) gossiping is conveyed by a verbal form for ‘tooth’ with -inė-. In line with Arad (2005:243), I assume that only a root level merge can yield such a range of idiosyncratic interpretations.

8.3 The aspectual effects of -inė- are epiphenomenal

In this section I argue that the effects of -inė- with regard to aspect are epiphenomenal and fall out as a result of its Number modifier status.

I draw parallels between -inė- and the aspectual effect of nominal Number (8.3.1). Then I show how the aspectual effects of -inė- can be predicted for each subclass of aspect (8.3.2).

8.3.1 Number: parallels between nominal and verbal domains

The importance of nominal arguments and their Number to the aspectual interpretation of the sentence has long been noted (Verkuyl 1973). For example, when the object is singular, the sentence is interpreted as telic or having an endpoint (24a); if the object is nonspecific plural, the sentence is interpreted as imperfective (24b):

(24) a. She ate an apple #?for an hour/in an hour.
   b. She ate apples for an hour/#in an hour.

The contrast in aspectual values is signaled by the compatibility with adverbial phrases ‘in x time’ and ‘for x time’ respectively. The aspectual effect notwithstanding,

---

8 Dr Jurgis Pakerys (p.c.) points out that tupinėti is also attested, but is not used in standard Lithuanian.
nominal Number is not considered an inherently aspectual morpheme. By analogy, verbal Number morphemes should not be labeled aspectual either.

I propose that regardless of the category domain—V or N—Number has the same aspectual effect. Namely, the effect of Number is a result of quantization (Krifka 1998). If within a VP there is a morpheme that cuts out a bounded portion of an event or individual, the entire VP is interpreted as quantized and therefore telic (25a–b). Conversely, if a VP contains a morpheme that is unbounded, the VP as a whole is interpreted as cumulative and therefore atelic (25c–d). Schematically, this could be illustrated as below:

(a) VP_{telic}  
\[ \text{V} \prec \text{DP}_{\text{singular}} \]
(b) VP_{telic}  
\[ \text{V}_{\text{singular}} \prec \text{DP} \]
(c) VP_{atelic}  
\[ \text{V} \prec \text{DP}_{\text{plural}} \]
(d) VP_{atelic}  
\[ \text{V}_{\text{plural}} \prec \text{DP} \]

The singular count nouns or events are quantized and therefore telic; the plural and mass nouns or events are cumulative and therefore atelic. Nominal effects have been known for some time, as discussed above. The attempt to provide a unified approach across the two domains is one contribution of this study. Specifically, I posit that the effect of -\text{inė}- is cumulative and thus atelic (25d). The proposal also correctly predicts the existence of quantizing verbal morphemes (25b) which I discuss elsewhere (Armoskaite and Sherkina-Lieber 2008).

8.3.2 Apectual effects of -\text{inė}-

I will now show the effect of -\text{inė}- on the four types of aspectual classes: activities (8.3.2.1), accomplishments (8.3.2.2), achievements (8.3.2.3) and statives (8.3.2.4).

8.3.2.1 The effect of -\text{inė}- on activity verbs  
Recall that activity predicates are defined as [−telic], [ + stage] (section 8.1.1). Given that -\text{inė}- encodes plural Number of events, its aspectual effect is cumulative, i.e. adding [ − telic]. Since activity verbs are already [ − telic], -\text{inė}- does not have any aspectual effect on activities:

(26)  
\[ \text{ūkinink-as pjaust-ė lap-us kiaurąq dieną} \]  
\[ \text{farmer-NOM.SG cut-PAST3SG leaf-ACC.PL whole day-ACC.SG} \]  
\[ \text{The farmer was cutting the leaves the whole day.} \]
b. Ŭkininkas pjaustinėjo lapus kiaurą dieną.
ūkinink-as pjaust-inė-jo lap-us kiaurą dien-q
farmer-NOM.SG cut-SUF-PAST3SG leaf-ACC.PL whole day-ACC.SG
‘The farmer was cutting the leaves the whole day (and took the task lightly).’

Thus, both (26a) and (26b) are atelic. However, there is a distinction between the two predicates: it is the difference in meaning due to the Number suffix. When -inė- is added to a verb like cut, the event of cutting is pluralized. The many instances of cutting render the verb diminutive and somewhat pejorative in the given context. If -inė- were a dedicated aspectual morpheme this contrast in meaning and lack of contrast in aspectual values would be hard to explain. Under the view that -inė- is Number, the lack of contrast in telicity is expected. Since the predicate is already atelic, -inė- only contributes to the interpretation of the number of the cutting events. The diminutive/pejorative shade of meaning of (26b) is a result of the interaction of -inė- with the lexical meaning of the verb (for the evaluative effects of -inė-, see some more discussion in the appendix). Note also, that given the root modifier status of -inė- (as discussed in sections 8.2.2–8.2.3.), this effect is expected, too.

8.3.2.2 The effect of -inė- on accomplishment verbs In section 1.1, we have defined accomplishment predicates as [+telic], [+stage]; atelicity judgements are illustrated here in (27a–b) with the help of adverbs.

(27) a. Dizaineris perstatė baldus per dešimt minučių.
dizainer-is perstat-ė bald-us per
designer-NOM.SG rearrange-PAST3SG furniture-ACC.PL in
dešimt minuč-ių
ten minute-GEN.PL
‘The designer re-arranged the furniture in ten minutes.’

b. *Dizaineris perstatė baldus kiaurą dieną.
dizainer-is perstat-ė bald-us kiaurą dien-q
designer-NOM.SG rearrange-PAST3SG furniture-ACC.PL whole day-ACC.SG
Intended: ‘The designer re-arranged the furniture the whole day.’

Adding -inė- changes the [+telic] value of accomplishments into [−telic], as adverbial tests indicate:

(28) a. *Dizaineris perstatinėjo baldus per dešimt minučių.
dizainer-is perstat-inė-jo bald-us per
designer-NOM.SG rearrange-SUF-PAST3SG furniture-ACC.PL in
dešimt minuč-ių
ten minute-GEN.PL
Intended: ‘The designer kept re-arranging the furniture in ten minutes.’
b. Dizaineris perstatėjo baldus kiaurą dieną.

dizainer-is perstat-inē-jo bald-us kiaurq dien-q

designer-NOM.SG rearrange-SUF-PAST3SG furniture-ACC.PL whole day-ACC.SG

‘The designer kept re-arranging the furniture the whole day.’

Not only does -inē- transform the telicity value of the verb, but it also adds the meaning of multiple events of re-arranging furniture.

8.3.2.3 The effect of -inē- on achievement verbs

Achievements have been defined as [+ telic], [− stage] in 8.1.1. These predicates denote instantaneous events. The adverbials that go with achievements capture the brevity of the event:

(29) Grandinės trūko akimirkšniu.

grandin-ēs trūk-o akimirkn-iu

chain-NOM.PL burst-SUF-PAST,3PL instant-INST.SG

‘The chains burst instantaneously.’

With -inē- the [+ telic] value of achievement changes into [− telic]. The event is still instantaneous, but now the predicate depicts a series of these instantaneous events as adverbial tests indicate:


grandin-ēs trūk-inē-jo akimirkšniu

chain-NOM.PL burst-SUF-PAST,3PL instant-INST.SG

Intended: ‘The chains kept bursting instantaneously.’

b. Grandinės trūkinėjo kiaurą dieną.

grandin-ēs trūk-inē-jo kiaurq dien-q

chain-NOM.PL burst-SUF-PAST,3PL whole day-ACC.SG

‘The chains kept bursting the whole day.’

8.3.2.4 The effect of -inē- on stative verbs

Stative predicates are [− telic], [− stage], as has been discussed in 8.1.1. The predicates are selected by adverbials that describe unbounded events as illustrated below:

(31) Tu mylėjai kiaurą dieną.

t. -o mylē-ji-s kiaurq dien-q

you-NOM.SG love-PAST2SG.-RECIP whole day-ACC.SG

‘You made love (to another person) the whole day.’

Given that the stative predicates are already atelic, -inē- does not contribute any aspectual change with respect to telicity. As with activities, which are also atelic, the main contribution of -inē- is that of plural Number of events:

9 As a native speaker, I find the stative root merged with -inē- ungrammatical and cannot pass any judgement on the use of these predicatives (also see Geniušienė 1997). However, the online dictionary (www.lkz.lt) lists a number of similar examples and other native speakers accept them, therefore I assume this to be a dialectal difference.
8.4 Conclusions and further questions

Based on the puzzling data with -iné-, I have shown that the suffix cannot be treated as a dedicated aspectual morpheme although it has an impact on aspectual interpretations of telicity (8.1.2). An alternative analysis has been offered: the behavior of the suffix is best accounted for when it is reanalyzed as an instance of Number for events (8.2.1). Moreover, the root modifier status of verbal Number is in line with other idiosyncratic characteristics—optionality, distributivity, range of meanings—that would not be expected if the suffix were genuinely aspectual (8.2.2–8.2.3). Drawing on parallels with the effects of nominal Number on aspect, I argued for a unified account of Number effects across the two domains (8.3). Assuming that Number effects on Aspect can be reduced to contrast in cumulativity versus quantization, I contend that both VP and DP can host Number. The evidence from Lithuanian makes a good case for the proposed approach. However, this is only the beginning of the study of Number syntax in relation to aspect.

To date, there exist few studies on verbal plurality from the typological perspective (Cusic 1981, Dressler 1968, Mithun 1988, Xrakovskij 1997). A small number of studies on the formal properties of verbal plurality are also available (Landman 1996, Lasersohn 1995, Kratzer 2008). There are still fewer studies where the formal and typological approaches intersect (Bar-el 2008, Yu 2003, van Geenhoven 2004), i.e. where the morphosyntactic properties of verbal plurality are discussed.

The questions to be answered are still basic. For example, it is not really known what the syntactic expressions of plurality are. A quick survey of Xrakovskij (1997) reveals that the prefixal position is rare, and the suffixal position appears to be more frequent. Yet another attested alternative is alteration of the root, like in Chechen:

(33)  d. uttu ‘to pour’ – dyttu ‘to pour repeatedly’
    loosu ‘to sift’ – lyssu ‘to sift repeatedly’
    saga ‘to shine’ – siega ‘to shine repeatedly’

[ (Yu 2003: 292–3) ]

Even these preliminary comparisons indicate that Number in verbal domain may have more than one syntactic position. While the Lithuanian suffix attaches to a verb just above the root level, Chechen Number must appear even lower given that it is encoded as a change of root vowel quality:
In contrast, a language like English possesses only phrasal expressions to convey plural and distributive events:

(35)  
   a. Casey and Stacey together did eleven liftings.  
   b. Two children lifted two boxes. (Kratzer 2008: 15)  
   c. She used to keep saying that. (overheard at a talk)

The emerging picture, where the devil is in the details, gives rise to more questions. Can all syntactic expressions of plurality be reduced to differences between heads and modifiers? How does their syntactic status correlate with respect to their impact on Aspect? What are possible/attested levels of merge, e.g. root, suffix, prefix, phrasal? Are there any language-specific or cross-linguistically universal restrictions on the expression and syntactic position of Number? What are restrictions on the expressions of Number with respect to categories, i.e. can Number span across the lexical domains of noun and verb (Bar-el 2008)? Further, the existence of plural marking predicts the existence of a singular event marker. It is attested in Russian, Lithuanian and Finnish, but it remains to be seen if it is cross-linguistically universal (Armoskaite and Sherkina-Lieber 2008; Armoskaite and Koskinen 2008). What other domains, besides aspectual, does verbal plurality interact with?

Appendix

Evaluative effects of -*inė*-: diminutive & pejorative

I Diminutive

Diminutive interpretation is a result of pluralization of an atomic event interacting with the lexical meaning of a given verb. The sub-events imposed on an atomic event
turns the event into a series of smaller events which is at odds with the lexical meaning:

(1) kalti kalinėti
   kal-ťi kal-îné-ťi
   hammer-INF hammer-SUF-INF
   ‘hammer’ ‘hammer repeatedly, lightly’

II Pejorative

Pejorative interpretation is a further re-interpretation of the diminutive interpretation. The attained diminutive activity is perceived as taken too lightly for what the activity should be, i.e. sub-events imposed on an atomic event are at odds with the inherent lexical meaning.

(2) paišyti paišinėti
    paiš-y-ťi paiš-îné-ťi
    draw-SUF-INF draw-SUF-INF
    ‘draw’ ‘draw repeatedly, in a lighthearted way’
Decomposing the mass/count distinction: Evidence from languages that lack it*

MARTINA WILTSCHKO

9.1 Introduction

In this chapter I wish to establish that the mass/count distinction is not universally grammaticized. In particular, I show that neither Blackfoot nor Halkomelem classify their nouns as mass or count. This does however not mean that speakers of these two languages cannot distinguish between substances and individuals, respectively. Rather, it is the grammar that does not care about this distinction. In this respect, Blackfoot and Halkomelem differ significantly from English and German, two Indo-European languages in which the mass/count distinction plays an important role for nominal classification. Adopting a generative point of view, according to which all languages share core properties (universal grammar), the observed difference in nominal classification requires explanation. What determines whether or not a language makes use of the mass/count distinction as a nominal classification strategy?

To capture this source of language variation, I propose that we need to distinguish between

i) ontological properties of nominals: whether nouns denote substances or individuals; and

ii) categorical properties of nominals: whether nouns are categorized as mass or count nouns.

* I would like to thank my language consultants, the late Halkomelem elders Rosaleen George and Dr Elizabeth Herrling as well as my Blackfoot teacher Beatrice Bullshields for sharing their language with me. Research for this paper was funded by a SSHRC Standard Research Grant (“Grammatical Categories and the Universal Functional Hierarchy”) awarded to the author.
Ontological properties are likely universally available. That is, I suspect that all languages have nominals that denote substance and nominals that denote individuals. However, these ontologically based properties are about classifying the world and not the nouns that serve to name the things in the world. By hypothesis, these properties are not visible for the grammar of a given language. To enter into grammatical relations, such as—for example—selection, the distinction must be grammaticized, which amounts to saying that there must be categorical properties that the grammar can see. It is these categorical properties, which are responsible for the distributional differences that distinguish mass nouns from count nouns (cf. Clausen et al. 2010 for a similar view). Crucially, on the basis of languages that lack it, I argue that the categorical properties associated with the mass/count distinction cannot be equated with nor can they be derived from ontological properties. The core goal of this chapter is thus to explore the source of these categorical properties, a topic, which has been much debated in the literature, and still is—as evidenced by the publication of this volume.

The remainder of the chapter is organized as follows. In section 9.2, I establish that the mass/count distinction is not universally associated with categorical properties that serve to classify nouns. I show that neither in Halkomelem nor in Blackfoot is the mass/count distinction grammaticized. In section 9.3, I explore the source of the categorical properties. In line with much current research, I assume that categorical properties are associated with functional categories (see for example Borer 2005). In particular, I propose that the categorical properties associated with the mass/count distinction are due to a binary feature \([ \pm \text{bounded}]\), which is associated with a functional category I identify as nominal inner aspect (Rijkhoff’s 1991 Seinsart). In section 9.4 I show that it is necessary to distinguish the feature responsible for the categorical distinction (i.e. \([ \pm \text{bounded}]\)) from the functional category that hosts inner aspect. In particular, I argue that Halkomelem lacks inner aspect and consequently does not make available the basis for classifying nominals as either mass or count nouns categorically. As for Blackfoot, I argue that it does utilize the functional category inner aspect, but that it differs from languages with a grammaticized mass/count distinction in the content of the feature that associates with it. Instead of \([ \pm \text{bounded}]\) Blackfoot uses \([ \pm \text{animate}]\) as a categorical classification device for nominals. Finally, in section 9.5 I conclude.

9.2 The mass/count distinction is not universally associated with categorical properties

The main purpose of this section is to establish that the grammatical properties associated with the mass/count distinction cannot be equated with their ontological properties. Rather I argue that we need to distinguish between ontological properties and categorical properties (see also Bosweld de Smet 1997, Joosten 2003). The former
are associated with nominal roots (henceforth $\sqrt{R}$-properties). These ontological properties are the properties of the things we find in the world and which nominal roots name. We can, for example, distinguish between substances and individuals. Crucially however, this ontological distinction is not categorical: there are things that cannot be uniquely classified as one or the other (for a more fine-grained distinction see Clausen et al. 2010). On the other hand, however, the mass/count distinction appears to be based on the grammatical properties of a nominal phrase (henceforth $f$-property). This distinction is categorical in that there are a number of morphosyntactic diagnostics that clearly divide nominal phrases as belonging to one or the other.

To account for the categorical properties, I assume that such properties are always tied to functional structure (see Borer 2005, Marantz 1997). In particular, I will argue that there is a functional category, which I label $F$ in (1), and which hosts a binary feature [$\pm$ bounded]. (I address the question regarding the identity of this category in section 9.3.)

(1) $[\pm \text{ bounded}]\ f$-properties [$\sqrt{R}$-properties]

If $F$ is [$+$ bounded], the result is a count noun. In contrast, if $F$ is [$-$ bounded], the result is a mass noun. In this chapter I am not concerned with the relation between $\sqrt{R}$-properties and $f$-properties (see Wierzbicka 1985). Suffice it to say that typically substance nouns are realized as mass nouns and individual denoting nouns are realized as count nouns. But this mapping does not proceed in a one-to-one fashion—a crucial argument for the division of labor between $\sqrt{R}$-properties and $f$-properties. What I am concerned with here is to demonstrate, on the basis of language variation, the necessity to distinguish between $\sqrt{R}$-properties and $f$-properties and to develop an analysis that adequately captures the observed variation. In particular, I will show that in Blackfoot and Halkomelem the mass/count distinction is not associated with any categorical $f$-properties. In other words, Blackfoot and Halkomelem lack a grammaticized mass/count distinction. Any properties that appear to be attributable to a mass/count distinction derive from $\sqrt{R}$-properties.

9.2.1 The mass/count distinction in English (a quick and simplified review)

To establish a baseline for the kinds of $f$-properties we may expect from a grammaticized mass/count distinction, I start with a brief overview of the ones identified for English (see Borer 2005 among many others). The first cluster of properties has to do with selectability. Some determiners and quantifiers are sensitive to the mass/count distinction as illustrated in (2–3).

(2) a. the/a/this/that/one/every/each/no tree [English]
b. these/those/two/several/some/many/no/all trees
c. *much/*little tree(s)
Second, mass nouns but not (singular) count nouns can function as bare arguments.

Third, only count nouns, but not mass nouns can be pluralized.

Fourth, only count nouns, but not mass nouns can be directly preceded by numerals.

Fifth, if a language has a grammaticized mass/count distinction for nouns it has strategies for reclassification. For example, English has classifiers naming the unit of natural occurrence of a substance and such classifiers turn mass nouns into count phrases.

Crucial evidence for decomposing the mass/count distinction into f- and \( \sqrt{R} \)-properties comes from the fact that the value of the grammaticized mass/count distinction is not always predictable from \( \sqrt{R} \) properties. If it were, one distinction would suffice. That is, the existence of form-meaning mismatches such as the ones listed in supports the dual source of the mass/count distinction.

Even though the classification of such nouns as mass or count is arbitrary, once they are classified as mass, grammar blindly treats them as mass nouns, despite the fact
that they denote individuals. Consequently, such nouns cannot be preceded by numerals (12), cannot be pluralized (13), can combine with determiners selecting for mass (14), and can function as bare arguments (15).

(12) a. *There is a furniture/silverware/grass/luggage in my garden.
    b. *There are five furnitures/silverwares/grasses/luggages in my garden.
    c. *There is five furniture/silverware/grass/luggage in my garden.

(13) a. There is furniture/silverware/grass/luggage/homework in my garden.
    b. *There are furnitures/silverwares/grasses/luggages in my garden.

(14) a. ?the/*a/this/that/*one/*every/no furniture/silverware/grass/luggage
    b. *these/*those/*two/*several/*some/*many furnitures/silverwares/
       grasses/luggages
    c. some/no/all/much/little furniture/silverware/grass/luggage

(15) I saw furniture/silverware/grass/luggage/homework

Thus, while ontologically these nouns would be classified as individual denoting (rather than substance denoting) nouns, for the purpose of the grammar they nevertheless behave like mass nouns. This supports the decomposition proposed in (1).

Further evidence for the independence of ontological distinctions from grammatically relevant categorical ones comes from the fact that ontological properties can be available even when the categorical ones are not. For example, in the context of root compounds, which by hypothesis lack functional structure, the difference between substance and individual roots is still available as shown in (16). Since only substances can be poured only substance denoting roots can function as the non-head part of the compound headed by pourer. In contrast, since only individuals can be picked, only individual denoting roots can function as the non-head part of the compound headed by picker.

(16) a. water pourer  #water picker
    water pourer  #wine picker
    wine pourer  #berry picker
    berry pourer  #flower pourer
    flower pourer  flower picker

Moreover, Harley (2005) shows that in the context of denominal verbs the mass/count distinction plays a role in determining the aspectual properties of the VP. Denominal verbs based on mass nouns yield atelic events (17), while denominal verbs based on count nouns yield telic events (18).

(17) a. The mare foaled  #for 2 hours/ in 2 hours.
    b. The dog whelped  #for 2 hours/ in 2 hours.
    c. The cow calved  #for 2 hours/ in 2 hours.
a. The baby drooled for 2 hours/# in 2 hours.
b. The athlete sweat for 2 hours/# in 2 hours.
c. The wound bled for 2 hours/# in 2 hours.

Assuming that such word formation is derived via root incorporation, it follows that the functional structure responsible for the grammaticized properties cannot be available. If so, this implies that even at the root level, there is a sense in which the mass/count distinction plays a role. In our terms, it is the \( \sqrt{R} \)-properties rather than the \( f \)-properties which are available. Consequently, we predict that in the case of mismatches between \( \sqrt{R} \)-properties and \( f \)-properties, it is the \( \sqrt{R} \)-properties that will win when the word occurs in environments that exclude \( f \)-properties. This is indeed the case, as shown in (19).

furniture picker
silverware picker
grass picker

Having established that the mass/count distinction familiar from languages like English has two independent sources, one ontological and one grammatical, we can now turn to the two lesser studied languages that form the empirical basis of our investigation.

9.2.2 The mass/count distinction in Halkomelem and Blackfoot

In this subsection, I demonstrate that the mass/count distinction is not universally grammaticized. In particular, I show that the relevant \( f \)-properties that categorize mass/count nouns as such are not available in Halkomelem (Central Coast Salish) or in Blackfoot (Algonquian).\(^1\) First, we observe that unlike in English, the mass/count distinction is not subject to selectability. To the best of my knowledge, there are no determiners or quantifiers that are sensitive to this distinction. This is illustrated in (20)–(21) for Halkomelem and in (22)–(23) for Blackfoot.\(^2\)

(20) a. tsel kw’êts-lexw qex (te) syîts’em/siyólh/qó/mélk. [Halkomelem]
1SG.S see-TRANS QU DET sand/wood/water/milk
‘I saw lots of sand/wood/water/milk.’
b. tsel kw’êts-lexw qex (te) theqát/sth’im/swíweles.
1SG.S see-TRANS QU DET tree/berry/boy
‘I saw lots of trees/berries/boys.’

1 Rhodes (1990) shows that Ojibwe, another Algonquian lacks a mass/count distinction. But see Mathieu (this volume) for a different view.
2 Unless noted otherwise, all data come from my own fieldwork.
a. *tsel kw’éts-lexw mekw’ (te) siyits’em/sqélep/siyólh. [Halkomelem]
   1sg.s see-trans qu det sand/dirt/wood
   ‘I’ve seen all the sand/dirt/wood.’

b. *tsel kw’éts-lexw mekw’ (te) theqát/sth’im/swíweles.
   1sg.s see-trans qu det.pl tree/berry/boy
   ‘I’ve seen all the trees/berries/boys.’

a. *nitohkanaissimatoop’ annihkayi aohkíí. [Blackfoot]
   nit-ohkan-a-simatoo’p anni-hka-yi aohkíí
   1-all-dur-drink.vti-2/1>in det-invis-in.sg water
   ‘I drank (up) all of that water.’

b. nitohkannainowayi anniksisk pookááiks.
   nit-ohkan-a-in-o-aa-yi ann-iksi pookáá-iksi
   1-all-dur-see.vta-dir-pl det-pl child-pl
   ‘I saw all the children.’

a. 1’i’násikóónsko  ii. 1’nákaohkíí  iii. 1’nákónnikis [Blackfoot]
   i’nák-kóónssko  i’nák-aohkíí  i’nák-ónnikis
   small-snow  small-water  small-milk/breast
   ‘a little bit of snow’  ‘a little bit of water’  ‘a little bit of milk’

b. 1’i’náksipokaa  ii. 1’nákánao’kssi  iii. 1’nákónnikis
   i’nák-pokaa  i’nák-ánao’kssi  i’nák-ónnikis
   small child  small-halfdollar  small-milk/breast
   ‘baby’  ‘quarter of a dollar’  ‘small breast’

Secondly, in neither one of the two languages does the mass/count distinction have
an effect on the syntax of arguments. That is, all nouns must be preceded by a
determiner, independent of whether they denote a substance or an individual. This is
shown for Halkomelem in (24) and for Blackfoot in (25).

a. *tsel kw’éts-lexw *(te) syits’em/sqélep/siyólh. [Halkomelem]
   1sg.s see-trans det sand/dirt/wood
   ‘I’ve seen sand/dirt/wood.’

b. í:mex *(te) swíyeqe.
   walking det man
   ‘The man is walking.’

a. *nitaissimatoop *(omi) aohkíí. [Blackfoot]
   nit-a-simatoo’p omi aohkíí
   1-dur-drink.vti-2/1>in dem water
   ‘I’m drinking that water.’
b. nitsinoaa *(oma) piita.
   nit-ino-aa oma piita
   1-sec.VTA-DIR DET eagle
   ‘I saw the eagle.’

Next, pluralization is not sensitive to a mass/count distinction: all nouns no matter whether they denote substances or individuals can be pluralized in both Halkomelem (26) and Blackfoot (27)

(26) **Singular** | **Plural** | [Halkomelem]
---|---|---
\(a\). th’exet | th’exth’éxet | gravel
\(b\). syiq | syiqyiq | snow
\(c\). spiw | spepiw | ice
\(d\). shwathetel | shweláthetel | fog
\(e\). speháls | spelháls | wind

(27) **Singular** | **Plural** | [Blackfoot]
---|---|---
\(a\). sopo | ikkináisopoistsi | ‘wind’
\(b\). aaapan | aaapaistsi | ‘blood’
\(c\). aiksinoosak | aiksinoosakiksi | ‘bacon’
\(d\). isstsskáán | isstsskäistsi | ‘dust’
\(e\). isttsiksipoko | isttsiksípokoistsi | ‘salt’
\(f\). kaatsi | káatsiistsi | ‘driftwood’
\(g\). kokóto | kokótoistsi | ‘ice’

(\(Frantz & Russell 1995\))

Finally, it is a striking fact about both Halkomelem and Blackfoot that there does not seem to be a dedicated strategy for reclassification. In Halkomelem, for example,
there are no classifiers that name the unit of natural occurrence of a substance such as *grain* or *drop*. Instead, there are other strategies to express those meanings. For example, there is a specialized form with the meaning ‘drop of water’ which is not transparently related to the form for ‘water’.

(28) a. th’q’ém/th’eq’ém  
    ‘drop of water’

b. qo  
    ‘water’

Moreover, regular modifiers can be used in contexts where in English we would use a classifier. Consider the examples in (29). The modifiers *i’axwíl* (‘small’) and *emémel* (‘small’) can be used with substance nouns to refer to individual grains. This simply reflects the ontological fact that a grain is a small amount of sand. Consequently, the same form can also be used to refer to any small amount of sand, not necessarily just a grain. We thus observe that there are two possible translations for this sentence.

(29) tsel kw’éts-l-exw *i’axwíl* syíts’em.  
    1sg.s see-trans-3o small sand  
    i) ‘I’ve seen a grain of sand.’  
    ii) ‘I’ve seen a little bit of sand.’

Crucially, the same two modifiers can also be combined with substance nouns that do not have a natural form of occurrence. In this case the result can be individuated (the ‘small piece’ reading) but it need not be (the ‘small amount’ reading).

(30) tsel kw’éts-l-exw (te) *i’axwíl* siyólh.  
    1sg.s see-trans-3o det small wood  
    i) ‘I saw a piece of wood.’  
    ii) ‘I saw a little bit of wood.’

And finally, the same modifier can also co-occur with nouns denoting individuals and in this case the smallness refers to the size of the individual.

(31) tsel kw’éts-l-exw *i’axwíl* theqát/ theqtheqát.  
    1sg.s see-trans-3o small tree/tree.pl  
    ‘I saw a small little tree/small little trees.’

This clearly establishes that these modifiers do not serve the function to turn a mass noun into a count phrase.

The same holds true for Blackfoot as well. We have already seen one such example involving the modifier *i’nák* in (23a/biii) above. Here the same noun can be used to refer to an individual, which is inherently bounded (i.e. ‘breast’) or a substance, which is not inherently bounded (i.e. ‘milk’). Crucially, in both cases *ónnikis* can be modified by *i’nák*. The same phenomenon can also be observed in the examples in (32).
On several occasions I have asked my consultant how to say things like ‘several pieces of wood’ or ‘snowflake’ and she would consistently use nouns with a modifier that is compatible with both substance and individual nouns. In fact, she told me that ‘We don’t get into this kind of stuff’.

This much establishes that Halkomelem and Blackfoot do not formally classify their nouns along a distinction in boundedness. However, this does not mean that nouns cannot refer to either substances or individuals, just that this ontological distinction does not map onto a formal classification system. There is however one property of substance nouns in both languages which may suggest that something like the mass/count distinction is nevertheless active. Sometimes, but crucially not always, substance nouns can be introduced by a numeral. As shown in (33), in Halkomelem sand and wood can be preceded by a numeral but snow and wind cannot. And as shown in (34), in Blackfoot wood and blood can be preceded by a numeral, but snow cannot.

(33) a. tsel kw’éts-l-exw isále siyítsem/ siyólh. [Halkomelem]
   1sg.s see-trans-3o two sand.pl/wood
   ‘I’ve seen two {pieces of, kinds of} sand/wood.’

b. *tsel kw’éts-l-exw isále syiqyiq/pehals
   1sg.s see-trans-3o two snow.pl/wind
   ‘I’ve seen two snow/wind.’

(34) a. nitoohkoonimh niisitoyi miistsís. [Blackfoot]
   nit-ohkooni-m-hp-yi niisitoyi miistsís
   1-find.vti-1>in-indep-sg five wood
   ‘I found five pieces of wood.’

b. naato’kayi a’apannists iihtsokina.
   nááto’k-ayi aáápan-ists iiht-sokin-aa
   two blood-pl inst-treat.vta-dir
   ‘There were 2 blood (bags) used to treat him.’

Consultant’s comments: you can say this in the hospital
I conclude that the possibility for a numeral preceding a noun (i.e. countability) is sensitive to ontological \( \sqrt{R} \)-properties rather than grammatical \( f \)-properties and consequently it does not behave in any systematic way. This is consistent with the fact that my Blackfoot consultant has given different grammaticality judgements for counting water on different occasions as shown in (35).

\[(35) \quad \text{a.}\ \text{*nitohkanaisimatoop niisitoyi aohkií. [Blackfoot]}
\]
\[
\text{nit-ohkan-a-simatoo-'}^p \niisitoyi\ aohkií
\]
\[
1\text{-all-DUR-drink.vti-2/in} \text{ five water}
\]
\`
I drank five water’.
\]
\[
\text{b.}\ \text{nááto’kayi aohkiísts}
\]
\[
\text{nááto’kayi aohkií-istsi}
\]
\[
two \text{ water-pl}
\]
\`
two water’
\]
Consultant’s comments: you can use this for two containers of water

9.2.3 Summary

In this section we have seen that the mass/count distinction manifests itself in different ways across different languages. This is summarized in Table 1.

Assuming a non-trivial distinction, between the ontological and the categorical properties underlying the mass/count distinction, we can understand the differences in behavior. Specifically, where English differs from Halkomelem and Blackfoot categorically, I assume that the source of variation lies in the \( f \)-properties. These are the properties that grammar cares about. In contrast, where judgements are not categorical but unstable, I assume that it is the \( \sqrt{R} \)-properties that are responsible. In this case grammar does not seem to care. This is summarized in (36).

**Table 9.1. Cross-linguistic differences in the manifestation of the mass/count distinction**

<table>
<thead>
<tr>
<th>Properties of the mass/count distinction</th>
<th>English</th>
<th>Halkomelem/Blackfoot</th>
</tr>
</thead>
<tbody>
<tr>
<td>Can be selected by determiners/quantifiers</td>
<td>yes</td>
<td>no</td>
</tr>
<tr>
<td>Can function as bare arguments</td>
<td>yes</td>
<td>no</td>
</tr>
<tr>
<td>Can be pluralized</td>
<td>no</td>
<td>yes</td>
</tr>
<tr>
<td>Can be introduced by numerals</td>
<td>no</td>
<td>sometimes</td>
</tr>
</tbody>
</table>
As we have seen, there are two reasons to separate ontological $\sqrt{R}$-properties from categorical $f$-properties. On the one hand, in English, a language with a grammaticized mass/count distinction, we observe certain mismatches between the ontological and the categorical properties. This suggests that the latter cannot be derived from the former. Moreover, in languages which lack a categorical distinction between mass and count nouns, we observe evidence for the ontological $\sqrt{R}$-properties but not for the categorical $f$-properties.

At this point, a brief discussion of the status of countability is in order. If the present proposal is on the right track, we have to conclude that whether or not a noun can be preceded by a numeral, i.e. whether or not it is countable, does not belong to the set of $f$-properties. In English, countability appears to be grammaticized and thus belong to the set of $f$-properties—hence the term ‘count noun’. That is, no mass noun can be preceded by a numeral. I argue however, that this pattern is misleading. Rather, the ability to be counted is an ontological property. However, in English the use of numerals interacts with the inflectional number system (i.e. the contrast between singular and plural); and this inflectional contrast is sensitive to the grammaticized mass/count distinction.

If this is the case, then there is an important lesson to learn here. Since $\sqrt{R}$-roots in Halkomelem and Blackfoot can be counted without the presence of a grammaticized mass/count distinction we can conclude that $\sqrt{R}$-roots do not require individuation or partitioning in order to interact with the count system (contra Borer 2005, Chierchia 1998b). If this is so, we predict that in English as well, roots should be countable even if the functional layer responsible for the mass/count distinction is absent. This prediction is indeed borne out. In the context of compounds, which lack the functional layer responsible for the mass/count distinction, numerals are possible independent of whether the nominal root denotes an individual as in (37) or a substance as in (38).

(36)  [$f$-properties  [$\sqrt{R}$-properties]]  
$f$-properties: categorical ('grammar cares')
$\sqrt{R}$-properties: ontological ('grammar doesn’t care')

(37)  a. three card trick, three colour process, three piece, three point turn [English]
b. four leaf clover, four letter word, four wheel drive, four stroke, four star
c. five finger exercise, five year plan
d. six day war

(38)  a. five spice, five spice powder
    b. five grain cereal

$^3$ Sometimes numerals are also found inside a phrasal compound, in which case we find plural marking on the counted noun: seven years war
Since countability is not dependent on the presence of a grammaticized mass/count distinction, it follows that it is not a reliable diagnostic for a grammaticized categorical mass/count distinction.

9.3 The source of the categorical properties of the mass/count distinction

We have now established that not all languages categorically classify their nouns as mass or count. I assume that the source of this variation is tied to a functional layer above the nominal root and not the root itself. This is consistent with the assumption that parametric variation is tied to functional categories (Borer 1983). The purpose of this section is to argue that the functional category responsible for formally classifying nominals as either mass or count is nominal inner aspect.

9.3.1 What is the categorial identity of F?

Thus far we have remained agnostic about the categorial identity of the functional layer that is responsible for the f-properties associated with the mass/count distinction. We have simply labeled it as F in (1), repeated below as (39) for convenience.

(39) \[ f_{F[±\text{bounded}]} f\text{-properties} [v_R v\text{-R-properties}] \]

So the question we are facing now concerns the identity of the functional category that hosts the features responsible for the mass/count distinction. Following Jackendoff (1991), I assume that the relevant feature is [± bounded]. In the verbal domain, this feature is associated with a functional category known as inner aspect (Travis 2010) found sandwiched in between the argument-structure defining verbal categories (v), as shown in (40a). Given the well-established parallel between verbal and nominal functional structure, I propose that inner aspect is a category also found in the nominal domain as illustrated in (40b). In English, verbal and nominal inner aspect both host the same feature, namely [± bounded].

(40) a. INFL
    \( \text{INFL} \rightarrow \text{Asp} \)
    \( \text{Asp} \rightarrow vP \)
    \( v \rightarrow \text{Asp} [±\text{bounded}] \)

b. D
    \( \text{D} \rightarrow \text{Num} \)
    \( \text{Num} \rightarrow nP \)
    \( n \rightarrow \text{Asp} [±\text{bounded}] \)
This proposal is a formal implementation of Rijkhoff’s 1991 proposal according to which the mass/count distinction reflects the nominal counterpart of verbal Aktionsart, namely Seinsart (see also Muromatsu 1998).

In what follows I eliminate two alternative analyses. Specifically, I show that neither the functional category NUMBER nor the semi-functional category \( n \) can be assumed to host the boundedness feature.

### 9.3.2 Is it NUMBER?

It has been argued that the functional category responsible for the mass/count distinction is the very same category that hosts number marking in languages like English and classifiers in languages like Chinese (Borer 2005, Chierchia 1998b, Ghomeshi 2003, inter alia). According to this view, roots are neither count nor mass. The addition of number marking in English yields a count noun, and according to Borer (2005), the absence of such marking yields a mass noun by default. This view is sketched in (41) below.

(41)\[
\begin{array}{c}
\text{D} \\
\text{D} \quad \text{NUMBER} \\
\text{NUMBER} \quad \ldots \\
\text{[+bounded: sg]} \\
\text{[–bounded: pl, mass]}
\end{array}
\]

The hypothesis that the relevant functional category responsible for number marking is also responsible for the grammaticized mass/count distinction appears to be supported by the Halkomelem facts. The absence of the grammaticized mass/count distinction correlates with the absence of grammaticized number in this language (Wiltschko 2005b, 2008). While in Halkomelem reduplication functions as a pluralizer, it is not associated with f-properties. As a consequence it is optional in the context of numerals and quantifiers (42). Moreover, plural marking does not trigger obligatory concord, as shown in (43).

(42)\[
a. \text{te lhíxw swíweles/ sóweles} \quad \text{[Halkomelem]}
\quad \text{DET three boy/ boy.pl}
\quad \text{‘the three boys’}

b. qex te s-th’im/ s-th’eth’im
\quad \text{many DET nom-berry/nom-berry.pl}
\quad \text{‘many berries’}
\]
On the view that the same functional category is responsible for the f-properties of number marking and for the f-properties of the mass/count distinction, the differences between English and Halkomelem can be straightforwardly accounted for in terms of variation in the availability of Number. In a language where Number is available, we expect to find f-properties to be associated with both the mass/count distinction and number marking. This is the case for English. In a language where Number is not available, we do not expect to find f-properties associated with either the mass/count distinction or plural marking. This is the case in Halkomelem.

However, the correlation between f-properties associated with number marking and the mass/count distinction breaks down in Blackfoot. In this language, the mass/count distinction is not associated with f-properties, but number marking is. This is summarized in Table 9.2.

### Table 9.2. Grammaticized mass/count vs. number marking

<table>
<thead>
<tr>
<th></th>
<th>English</th>
<th>Halkomelem</th>
<th>Blackfoot</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mass/count distinction has f-properties</td>
<td>✓</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Number marking has f-properties</td>
<td>✓</td>
<td>x</td>
<td>✓</td>
</tr>
</tbody>
</table>

Evidence that plural marking in Blackfoot displays f-properties comes from the fact that it triggers obligatory concord, and it is obligatory in the context of numerals.

As we have seen in section 9.2, however, plural marking is not sensitive to the mass/count distinction. All nouns in Blackfoot can be pluralized independent of whether they denote individuals or substances (see (27)).

The existence of the Blackfoot pattern thus establishes that we cannot tie the presence or absence of a grammaticized mass/count distinction to the presence or
absence of grammaticized number marking. This undermines the hypothesis that the functional category relevant for the f-properties of the mass/count distinction is \textit{Number}.

9.3.3 \textit{Is it GENDER?}

Another potential hypothesis regarding the locus of the mass/count distinction in the nominal domain would be to associate it with \textit{n}. This would amount to saying that the count/mass distinction essentially functions as a nominal classification device akin to gender in German. This possibility is sketched in (45).

\begin{center}
(45)
\[
D \quad n \\
\downarrow \quad \downarrow \\
D \ldots \quad n \quad \sqrt{\text{nominal root}}
\end{center}

\textit{[masc/fem/neut]} \textit{[+-bounded]}

Initial evidence for this possibility comes from the behaviour of diminutive suffixes. They are associated with neuter gender and they appear to turn mass/nouns into count nouns (van Riemsdijk, p.c. reported in Borer 2005: p. 92, fn.6; cf. Wiltshko 2007). This is illustrated in (46). The diminutivized form of the mass nouns \textit{Brot} and \textit{Schnaps} function as count nouns. This can be seen on the basis of the fact that the preceding quantifier shows plural agreement.

\begin{center}
(46) a. viel Brot a'. viele Bröt-chen [German] \\
\text{much bread} \quad \text{many.PL} \quad \text{bread-DIM} \\
\text{‘much bread’} \quad \text{‘many little sandwiches’} \\
b. viel Schnaps b'. viele Schnaps-erl \\
\text{much Schnaps} \quad \text{many.PL} \quad \text{Schnaps-DIM} \\
\text{‘much Schnaps’} \quad \text{‘many little (good) Schnaps’}
\end{center}

The mass/count distinction however co-exists with the gender distinction and crucially cuts across the values for \textit{Gender}. In particular, there are mass nouns of all three genders (47) just like there are count nouns of all three genders (48).

\begin{center}
(47) a. Ich habe den Wein getrunken. [German] \\
\text{I aux det.MASC wine drunk} \\
\text{‘I have drunk (the) wine.’} \\
b. Ich habe das Wasser getrunken. \\
\text{I aux det.NEUT water drunk} \\
\text{‘I have drunk (the) water.’}
\end{center}
c. Ich habe die Suppe getrunken.  
I aux det.fem soup drunk  
‘I have drunk (the) soup.’

(48) a. Ich habe den Apfel gegessen.  [German]  
I aux det.masc apple eaten  
‘I have eaten the apple.’

b. Ich habe das Semmerl gegessen.  
I aux det.neut roll eaten  
‘I have eaten the roll.’

c. Ich habe die Karotte gegessen.  
I aux det.fem carrot eaten  
‘I have eaten the carrot.’

In sum, the patterns of variation we observe in the occurrence of nominal classification devices suggests that a grammaticized mass/count distinction does not correlate with the presence of grammaticized Number or Gender as summarized in Table 9.3.

<table>
<thead>
<tr>
<th>Table 9.3. Typology of nominal classification devices</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>Engl</td>
</tr>
<tr>
<td>-----------------------------------------------</td>
</tr>
<tr>
<td>Mass count distinction has f-properties</td>
</tr>
<tr>
<td>Number has f-properties</td>
</tr>
<tr>
<td>Gender has f-properties</td>
</tr>
</tbody>
</table>

We conclude that the mass/count distinction is neither associated with Number nor with n. It follows that it must be associated with another functional category. By elimination, this supports Rijkhoff’s proposal according to which the mass/count distinction instantiates nominal inner aspect.

9.4 Variation in the content of the categorical properties:  
[ ± bounded] versus [ ± animate]

Thus far, we have established that the mass/count distinction is not universally grammaticized. Neither Halkomelem nor Blackfoot makes use of such a distinction. We take this to mean that in these languages the [ ± bounded] feature is not associated with inner aspect. There is however another question that arises in light of this analysis. Is it possible to have inner aspect in the absence of [ ± bounded]? Or does the lack of the feature [ ± bounded] necessarily imply the lack of the functional category that hosts it?
In this section I argue that both options are attested. Halkomelem lacks the functional category inner Aspect, as in (49a).\(^4\) Blackfoot has inner aspect, but it serves as the host for a different feature, namely \([\pm\text{animate}]\), as in (49b).

\[(49)\quad \text{a. Halkomelem} \quad \text{b. Blackfoot}\]

\[
\begin{array}{c}
\text{n} \\
\text{n} \\
\text{\&N} \\
\text{Asp} \\
\text{\[\pm\text{animate}\]}
\end{array}
\]

The assumption that Halkomelem lacks inner aspect captures the fact that there is no nominal classification device akin to the mass/count distinction. Consequently, no classificatory properties are listed in Galloway’s dictionary of the language. This contrasts with English, where dictionary entries for nouns list their classification as either mass or count. I have nothing else to add to the Halkomelem pattern. Instead I focus on Blackfoot and I argue that instead of the \([\pm\text{bounded}]\) feature, inner aspect hosts \([\pm\text{animate}]\). This captures the fact that the dictionary entries of all Blackfoot nouns list the value of this feature (Frantz and Russell 1995). I develop the argument as follows. I first argue that Blackfoot animacy is formally different from gender marking of the German type (9.4.1). I then show that it behaves formally similar to the mass/count distinction in English (9.4.2).

9.4.1 Blackfoot animacy is not a form of gender

According to the traditional Algonquianist view, animacy is a form of gender marking (Dahlstrom 1995, Darnell and Vanek 1976, Goddard 2002, Greenberg 1954, Hockett 1966, Joseph 1979). A contrastive examination of German gender and Blackfoot animacy, however, reveals that the two classification devices differ in formal and functional properties. For example, Kilarski 2007: 334 points out that the principal differences between Algonquian and Indo-European gender, apart from the different number of genders—usually two or three in Indo-European—involve the type of

\(^4\) That this is an option made available by UG is argued for in MacDonald (2008), who analyzes Russian as lacking inner aspect in the verbal domain.
assignment criteria: in contrast to Algonquian, semantic criteria in Indo-European are usually weaker, being combined with formal ones (morphological or phonological). Furthermore, sex, rather than animacy, is the primary distinction [...].

Here I am mainly concerned with the formal differences suggesting that we are dealing with two distinct nominal classification devices. I present two pieces of evidence. First, in German, all nominalizing suffixes are classified for gender; this is not true for Blackfoot animacy (9.4.1.1). Second, in German all nouns are associated with a unique value for gender; in contrast, there are numerous Blackfoot nouns that can be associated with both values for animacy (9.4.1.2).

9.4.1.1 Classification of nominalizing suffixes Nominalizing suffixes in German determine the gender of the noun (Köpcke 1982, Zubin and Köpcke 1984, Eisenberg and Sayatz 2004). For example, the suffix -ik attaches to roots, which do not exist as independent words, and derives feminine nouns (50). The suffix -er attaches to roots and derives masculine nouns (51).

(50) -ik → [fem]  
\[
\begin{array}{lll}
\text{a. die Grammat-ik} & \text{b. die Graf-ik} & \text{c. die Mus-ik} \\
\text{DET.F grammar} & \text{DET.F graphic} & \text{DET.F music} \\
\text{‘the grammar’} & \text{‘the graphic’} & \text{‘the music’}
\end{array}
\]

(51) -er → [masc]  
\[
\begin{array}{lll}
\text{a. der Lehr-er} & \text{b. der Fahr-er} & \text{c. der Gärtn-er} \\
\text{DET.M teacher} & \text{DET.M driv-er} & \text{DET.M garden-er} \\
\text{‘the teacher’} & \text{‘the driver’} & \text{‘the gardener’}
\end{array}
\]

There is evidence that the gender of the resulting noun is in fact dependent on the suffix rather than being determined by the root. Some nominalizing suffixes attach to existing nouns (as apposed to roots) which are already associated with gender. When suffixed with the nominalizer -in the resulting noun is of a different gender as shown in (52). This suggests that it is the suffix itself, which determines the gender of the newly derived noun.

(52) -in[masc] → [fem]  
\[
\begin{array}{lll}
\text{a. der Architekt} & \text{die Architekt-in} \\
\text{DET.M architect} & \text{DET.F architect-F} \\
\text{‘the architect’} & \text{‘the female architect’} \\
\text{b. der Lehrer} & \text{die Lehrer-in} \\
\text{DET.M teacher} & \text{DET.F teacher-F} \\
\text{‘the teacher’} & \text{‘the female teacher’} \\
\text{c. der Student} & \text{die Student-in} \\
\text{DET.M student} & \text{DET.F student-F} \\
\text{‘the student’} & \text{‘the female student’}
\end{array}
\]
The pattern in (50)–(52) suggests that nominal suffixes are associated with gender, and to the best of my knowledge, this is the case for all such suffixes. Note that this pattern also suggests that gender is associated with the lowest layer of nominal classification, as in (53).

(53) \[ \text{AspP Asp[ } \pm \text{ bounded] [nP n[gender] [\sqrt{\text{root}}]]} \]

If gender was associated with the higher position, we may expect nouns and nominalizing suffixes that do not uniquely determine the gender of a noun. This is precisely the pattern we observe in Blackfoot, as I will now show.

The Blackfoot nominalizing suffix *a’tsis* can derive [+ animate] nouns as in (54) as well as [− animate] nouns as in (55). Since in Blackfoot, plural marking varies with the value for animacy I use it as a diagnostic throughout this chapter.

(54) *a’tsis* → [+ animate] [Blackfoot]
   a. saa’kssoyaa’a’tsis saa’kssóyaa’tsiiksi
      ‘poison ivy’
      ‘poison ivy plants’
      (Frantz and Russell 1995: 151 [F&R 1995])
   b. aawápspiinao’sa’tsis sikawapspiina’sa’tsiiksi
      aawápspiinao’s-a’tsis sikawapspiina’s-a’tsis-iksi
      ‘eye-glasses’
      ‘black eye-glasses’
      (F&R 1995: 4)

(55) *a’tsis* → [− animate] [Blackfoot]
   a. isoohkama’a’tsis poksisoohkama’a’tsiistsi
      isoohkama-a’tsis poks-isooohkama-a’tsis-istsi
      ‘container’
      ‘little storage bags’
      (F&R 1995: 69)
   b. issáana’kima’a’tsis issáana’kima’tsiistsi
      issáana’kim-a’a’tsis issáana’kim-a’tsis-istsi
      ‘candle, lit. fat lamp’
      ‘candles’
      (F&R 1995: 69)

For completeness, note that it is not the underlying form which determines the gender of the derived form. The suffixation of *a’tsis* to a [+ animate] noun may result in a [− animate] noun, as shown in (56).

(56) *a’tsis* [+ anim] → [− anim] [Blackfoot]
   a. issitsímaaan issitsímaanikksi
      issitsímaan issitsímaan-iksi
      ‘baby’
      ‘babies’
      (F&R 1995: 70)
   b. issitsímaa’tsis nitsssitsíma’a’tsiistsi
      issitsíma-a’tsis nits-issitsíma-a’tsi-istsi
      ‘baby thing’
      ‘my baby things’
      (F&R 1995: 70)
We can understand the lack of animacy specification of *a’tsis* if we assume that this nominalizer is associated with the lower nominal layer and that animacy is a higher nominal classification device as shown in (57).

\[(57)\] 
\[
\text{AspP} \quad \text{Asp[ ± animate]} \quad \text{[nP n[-a’tsis] [ . . . ]]}\]

### 9.4.1.2 Classification of nouns

The second argument that animacy in Blackfoot is formally distinct from German gender stems from the fact that some nouns are associated with two values, with a different—albeit related—meaning. Consider the example in (58). The same form *miistís* can be used as a [−animate] noun to mean *stick or branch* or as a [+animate] noun to mean *tree*.

\[(58)\] 
\[
a. \quad [\text{−animate}] \quad 
\begin{array}{ll}
\text{miistís} & \text{miistísístsi} \\
\text{branch} & \text{branch-PL.INANIM} \\
\end{array} \\
\text{‘stick, branch’ ‘branches’} \quad \text{[Blackfoot]} \quad \text{(F&R 1995: 98)}
\]
\[
b. \quad [+animate] \quad 
\begin{array}{ll}
\text{miistís} & \text{miistísíksi} \\
\text{tree} & \text{tree-PL.ANIM} \\
\end{array} \\
\text{‘tree’ ‘trees’} \quad \text{(F&R 1995: 98)}
\]

This particular example is famous for it shows that animacy is not a notion that necessarily depends on the ontological properties of the referent. However, the general pattern appears to be moderately productive. That is, in Frantz and Russell’s 1995 dictionary there are several such examples.

What is of interest in the present context is that the Blackfoot animacy specification differs from the German gender specification in precisely this respect. There are no German nouns that are associated with two distinct genders and still related in meaning.\(^5\) If a given form has two possible genders associated with it, it is for one of the following two reasons. We are either dealing with accidental homophony or else gender is in free variation and does not correlate with a meaning difference.

I conclude that Blackfoot animacy is not a form of gender. It displays formal properties different from those associated with German gender: not all nominal suffixes are classified for animacy, and some nouns can be associated with two values. This is summarized in Table 9.4 below.

### 9.4.2 Blackfoot animacy is like the boundedness distinction

In this subsection I show that Blackfoot animacy formally behaves like the boundedness distinction, which gives rise to the categorical mass/count distinction in Indo-European languages.

\(^5\) The so-called common gender of Russian differs in this respect (see Steriopolo 2008, Steriopolo and Wiltschko 2010, for discussion).
In contrast to gender, the mass/count distinction is not always uniquely determined for a given noun. Consider the German nouns in (59)–(60). They can all be used as mass nouns, in which case they denote an unbounded substance, as in (59). When pluralized, these nouns must be interpreted as denoting bounded individuals, as in (60). In all these cases the bounded form is the special form in that its meaning is not completely compositional: the bounded form of water for example can be used to denote the little bottle of liquids used in a salon or spa (which may not even contain water). The bounded form of bread is used for sandwiches and the bounded form of light can be used for Christmas lights.

(59) [−bounded] a. viel Wasser ‘much water’ b. viel Brot ‘much bread’ c. viel Licht ‘much light’

(60) [+bounded] a. viele Wässer ‘many waters’ (i.e. in a hairsalon) b. viele Brote ‘many breads’ (i.e. sandwiches) c. viele Lichter ‘many lights’ (i.e. christmas lights)

This pattern is reminiscent of a pattern we find associated with Blackfoot animacy marking on nouns. There are many cases where the [−animate] form denotes the general referent, while the [+animate] is the special form. In this case it often names a culturally more recent item.


(62) a. [−animate] ko’s ‘dish (earthenware)’ b. [+animate] ko’s ‘dishes’


While I have nothing to say about the mechanism that underlies this pattern, it is of interest in the present context that animacy marking behaves like the boundedness distinction in German and not like the gender distinction. In this context, an example from Fox cited in Goddard (2002) is of interest.

\[(64)\]
\[
\begin{align*}
\text{a. } & \text{[–animate] } \text{ksiisą́́ki’taan } \text{ómahksksiisá́́ki’taanistsi} \\
& \text{‘arrowhead’} \text{ ‘arrowheads’} \\
\text{b. } & \text{[+animate] } \text{ksiisą́́ki’taan } \text{ksiisą́́ki’taaniksi} \\
& \text{‘cartridge’} \text{ ‘cartridges’} \quad \text{(F&R 1995: 94)}
\end{align*}
\]

In (64), the [–animate] form refers to a substance, while the [+animate] form refers to the bounded version of the substance. As such, animacy marking seems to play the role of individuation (Mathieu, this volume). This is consistent with the claim that it occupies the same position as the boundedness distinction in Indo-European languages.

A second piece of evidence for animacy patterning with the boundedness distinction stems from the fact that animacy marking is subject to selectability. As mentioned above, Blackfoot singular and plural marking is sensitive to animacy marking. As illustrated in Table 9.5, -wa and -íksi are singular and plural markings associated with [+animate] nouns while -yí and -ístsi are associated with [–animate] nouns.

I interpret the sensitivity of number marking to animacy as an indication of selectability, one of the formal diagnostics for the boundedness distinction.

Another diagnostic we have identified in section 9.2 concerned mismatches between the meaning of the root and the nominal classification associated with it. That is, we have seen that the value of the boundedness distinction is not always predictable from ontological properties. The same holds for the animacy distinction in Blackfoot. It cannot always be predicted on the basis of the ontological properties of the root.\(^6\)

<table>
<thead>
<tr>
<th>Table 9.5. Number marking is sensitive to animacy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Singl</td>
</tr>
<tr>
<td>[–animate]</td>
</tr>
<tr>
<td>[+animate]</td>
</tr>
</tbody>
</table>

\(^6\) This is a pervasive property of animacy marking across the Algonquian language family and has attracted much attention in the literature. In particular, it has served as a major argument in the claim that animacy is a formal property (Bloomfield 1933, Black 1969, Dahlstrom 1995, Darnell and Vanek 1976, Goddard 2002, Greenberg 1954, Lehmann 1958) though attempts have been made to at least account for (if not predict) these apparent mismatches in semantic terms (Hallowell 1960, Black-Rogers 1982).
We have already seen instances of this in (58) and (61)–(63). But such mismatches between ontological and grammatical properties are not restricted to nouns associated with both values. There are also ontologically inanimate nouns that are grammatically classified as [+animate] as in (65). These nouns do not have a corresponding [−animate] noun.

(65) a. pokón ‘ball’ d. moápssp ‘eye’ [Blackfoot]
    b. issk ‘pail’ e. naató’si ‘sun’
    c. isttoán ‘knife’ f. ksisís ‘thorn’ (F&R: 1995)

A final way in which animacy behaves formally like the boundedness distinction in Indo-European concerns its interaction with verbal aspect. It is well known that in English the boundedness distinction interacts with the verb to determine the telicity of the resulting VP. While an unbounded object (either mass or bare plural) derives an atelic VP, a bounded one derives a telic VP.

(66) a. Yesterday’s sun melted a snowflake (#but there is still some left).
    b. Yesterday’s sun melted snow (but there is still some left).
    c. Yesterday’s sun melted snowflakes (but there are still some left).

On the present view, the interaction between the mass/count distinction and verbal telicity reflects the fact that the same feature ([± bounded]) is associated with inner aspect in the verbal and in the nominal domain. The correlate of the former is telicity, the correlate of the latter is the mass/count distinction. In contrast, in Blackfoot, it is the animacy distinction that interacts with the classification of verbal phrases. Like other Algonquian languages, Blackfoot verb stems are sensitive to the animacy of the object (in case of transitive verbs) or to the animacy of the subject (in case of intransitive verbs). This suggests that in Blackfoot [+animate] substantiates not only nominal inner aspect but also verbal inner aspect. On independent grounds, Ritter & Rosen (2010) have argued that this is in fact the case.

I am not aware of any such interaction between gender and verbal inner aspect. Again, this makes Blackfoot animacy formally more similar to the boundedness distinction than to a gender distinction. I thus conclude that Blackfoot animacy marking is the formal equivalent of the mass/count distinction and as such is associated with inner aspect. This concludes our investigation of the range of variation associated with nominal inner aspect.

9.5 Conclusion

In this chapter, I have argued that languages display differences in the manifestation of the mass/count distinction. In particular, I have shown that neither in Halkomelem, nor in Blackfoot is the mass/count distinction associated with grammatical
f-properties. This contrasts with English where this distinction is in fact categorical. Speakers of Halkomelem and Blackfoot can nevertheless distinguish between substance and individual denoting nouns. I have argued that this reflects the ontological properties of the nominal roots, and that this is also available in English. But in order to detect it in English one has to investigate structures that lack the functional layer responsible for the mass/count distinction. Such structures are available in root compounds as well as in denominal verbs. In all other environments the grammatical classification of nominals as mass or count is obligatory. The diagnostics we have used to establish whether or not there is such a grammaticized mass/count distinction are summarized in Table 9.6.

Importantly, our study suggests that countability does not serve to diagnose the mass/count distinction. Instead it is only sensitive to the ontological properties of the nominal root. An important implication of this finding is that nominal roots do not require functional structure to be individuated in order to interact with the count system.

The empirical observation that the mass/count distinction is not a universal nominal classification device has led us to investigate its source as well as the range of variation associated with it. In particular, I have argued that the mass/count distinction is best analyzed as resulting from the feature [+ bounded] associated with nominal inner aspect. Neither Halkomelem nor Blackfoot make use of this feature to classify their nouns. However, there is more to language variation than simply the absence or presence of a specific feature, such as [+ bounded]. In particular, I have argued that in Halkomelem the absence of this feature reflects the absence of the functional category it associates with in English, i.e. inner aspect. In contrast, in Blackfoot, inner aspect is available, but it is associated with a different feature, namely [+ animate]. Consequently, in Blackfoot all nominals are classified as either animate or inanimate, and this classification is not fully determined by the ontological properties of the nominal root. In this respect animacy in Blackfoot is the formal and functional equivalent of the mass/count distinction in English.

### Table 9.6. Diagnosing the grammaticized mass/count distinction.

<table>
<thead>
<tr>
<th></th>
<th>English</th>
<th>Halkomelem</th>
<th>Blackfoot</th>
</tr>
</thead>
<tbody>
<tr>
<td>Can be selected by determiners/quantifiers</td>
<td>yes</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>Can function as bare arguments</td>
<td>yes</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>Can be pluralized</td>
<td>no</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>Can be preceded by numeral (countability)</td>
<td>no</td>
<td>sometimes</td>
<td>sometimes</td>
</tr>
</tbody>
</table>

Root properties
This supports the view advocated in Ritter & Wiltschko (2009) according to which language variation may result from different features substantiating the same functional category (i.e. the parametric substantiation hypothesis). It also suggests that functional categories are not merely defined by the features that comprise them, but instead that there is a universal hierarchy of functional categories available, independent of the features that associate with them.
On the mass/count distinction in Ojibwe

ERIC MATHIEU

10.1 Introduction

On the basis of the fact that Ojibwe allows the pluralization of many nouns, including mass nouns, it has been suggested in the literature that Ojibwe lacks a grammaticized mass/count distinction. Consider, for example, the quotes in (1) and (2) taken from the functionalist literature.

(1) ‘Or consider the Algonquian language Ojibway (Richard Rhodes 1990:153–4, and personal communications). Nouns which might be expected not to have a plural do in fact form plurals freely, interestingly with the unit reading and not with the sort reading. Thus mkwam ‘ice’ or ‘piece of ice’, mkwamiig (plural) ‘pieces of ice’. Rhodes is unable to find a noun that cannot be pluralized in Ojibway.’ (Corbett 2000:87)

(2) ‘In Ojibwa there is no grammatical distinction like the mass/count distinction of Indo-European. Thus mkwam can equally mean ‘ice’ or ‘a piece of ice’. Nbiish can mean ‘water’ or ‘an amount of water.’ (Rhodes 1990:153)

The idea is that while the count/mass distinction has a clear morphosyntactic effect in English, this is apparently not the case in Ojibwe. As is well known, count nouns in English exhibit a singular/plural contrast nut (singular) versus nuts (plural), whereas mass nouns do not. Although grass is possible, the plural version of grass is not: *grasses (abstracting away from kind readings). Ojibwe certainly behaves differently from English in that regard: as pointed out in (1) and (2), one can pluralize, not only count nouns such as baagan ‘nut’ ~ ‘baaganan’ ‘nuts’, but also mass nouns (on a non-kind reading). The following examples, whose English translation both necessarily involves a measure phrase (the noun is not interpreted

1 As is well known, pluralized mass nouns in English can have a kind reading (see Chierchia 1998b and many others).
as a kind), illustrate the phenomenon in the context of a sentence.\(^2\) The object noun agrees with the verb in number and gender (Ojibwe, like other Algonquian languages, has a gender system based on animacy).

\(\begin{align*}
(a) & \quad \text{n-gii-waabam-aa-g mandaamin-ag.} \\
& \quad \text{1sg.s-past-see-3o-pl.anim corn-pl.anim} \\
& \quad 'I saw pieces of corn.' (literally: 'I saw corns.') \\
(b) & \quad \text{n-gii-waabam-aa-g mashkosiw-ag.} \\
& \quad \text{1sg.s-past-see-3o-pl.anim grass-pl.anim} \\
& \quad 'I saw blades of grass.', literally: 'I saw grasses.'
\end{align*}\)

The general ability of nominals to be pluralized in Ojibwe resembles the case of Halkomelem Salish, a language which, like many other Native North American languages (Mithun 1988), has been argued to lack a grammaticized mass/count distinction (Wiltschko 2007, 2008). In that language, all nouns appear to be the target of pluralization, including mass nouns (see also Davis and Matthews 1999 for Lillooet Salish). According to Wiltschko’s (2007, 2008) account, this follows from the fact that, whereas number marking in English is a functional head (#, head of #P), in Halkomelem Salish it is a modifier (to the nominal root). As a modifier, plural marking can attach to categories other than count nouns, including mass nouns.

Although Halkomelem Salish and Ojibwe share the property of systematic noun pluralization, I want to argue in this chapter that this property arises independently from the parameter setting proposed by Wiltschko (2007, 2008). I show in particular that pluralization of mass nouns is possible in languages where number marking is inflectional. This is true for some Indo-European languages, which means that Halkomelem Salish may not in this regard be very different after all from these better-studied languages. Pluralized mass nouns in Halkomelem Salish appear to yield a reading that is similar to an ‘abundance reading’ (to use a term introduced by Corbett 2000), or at least a reading where the exact counting of individuals is neither possible or important, and which we find in languages such as French, Hebrew, Persian and even (Biblical) English.

On the other hand, pluralization of mass nouns in Ojibwe yields a unit or a unit of measure reading with a clear atomization effect. This, I will argue, is made possible in Ojibwe, because its grammar embeds a singulative number system. I will show that, as in other singulative languages, the singulative in Ojibwe is expressed morpho-syntac-

\(^2\) Many of the data used in this article are taken from fieldwork undertaken with members of The Chippewas of Nawash Unceded First Nation at Cape Croker (Neyaashiinigmiing). I wish to thank Philomene Chegahno, Berdina Johnston, Donald Keeshig, Joanne Keeshig, Isabel Millette, Juanita Pheasant, Ernestine Proulx, and Ella Waukey for teaching me Ojibwe. Thanks also to Shirley Ida Williams from Trent University for sharing her knowledge with me. Thanks are also due to Richard Rhodes, Bethany Lochbühler, Elizabeth Ritter, Hagit Borer, Jila Ghomeshi, and Martina Wiltschko for their questions and comments. Funding by SSHRC is gratefully acknowledged: #230424-120699-2001 and #230611-120699-2001.
tically through gender shift. While in Breton, Arabic, etc. gender shift is from masculine to feminine, in Ojibwe it is from inanimate to animate. Although this process is admittedly less than transparent in Ojibwe, because of the loss of final vowels on singular nouns (final -a for animate nouns, final -i for inanimate nouns), there are remaining exemplars nonetheless. I will show that Fox (Mesquakie), another Algonquian language, provides clear evidence that Algonquian grammars embed a singulative system, since in that language final -a and -i on nouns have been retained.

In the course of the discussion, I will also review Wiltschko’s (this volume) newer account of the count/mass distinction in Halkomelem Salish and in particular her proposal, based on Blackfoot, that animacy might not be gender, but nominal aspect. While her claims may well be correct for Blackfoot, in Ojibwe there is every reason to believe, as I will show, that animacy is gender in Ojibwe and not nominal aspect.

Section 10.2 shows that number in Ojibwe is clearly inflectional. Section 10.3 introduces the relevant examples showing that plural mass nouns as well as singular mass nouns can be used in Ojibwe to express units from collectives and units of measure for mass nouns. Section 10.4 gives an analysis of these facts and provides an account of the singulative within Borer’s (2005) theory of number and division. Section 10.5 concludes the chapter.

10.2 Number as an inflectional category in Ojibwe

Many Native North American languages have been claimed to lack inflectional plural marking and more generally the traditional grammaticized mass/count distinction found in languages of the Indo-European type. Consider the following quote from Mithun (1988).

(4) ‘In the vast majority of North American languages … only certain nouns have plural forms. In most of these, only nouns referring to human beings have plurals, or only some nouns referring to humans, often kin terms. (Multiple animals that are considered ‘sentient things’, such as pets or characters in legends, are also often referred to by plural nouns.) The plurals that do exist are used only on some occasions, not every time multiple participants are discussed.’ (Mithun 1988:212)

For example, in Slave, an Athabaskan language spoken in parts of the Northern Territories, British Columbia and Alberta, Canada, a plural suffix -ke is attached optionally to nouns denoting a human or a dog, but never to inanimate nouns (Rice 1989). In other Native North American languages plural marking is completely optional and plural marking can target all sorts of nouns, count or mass. This is the case in Halkomelem Salish (Wiltschko 2008), and according to Wiltschko this suggests that the mass/count distinction is not grammaticized in that language. Hopi

3 ‘In languages with such systems, number can usually be determined from the verb.’ Mithun (1999:83).
(Whorf 1941) and Lillooet Salish (Davis and Matthewson 1999) are other North American languages for which the grammaticized mass/count distinction has been claimed to be lacking.

Although the lack of a grammaticized mass/count distinction appears widespread in Native North American languages, there are notable exceptions. Mithun (1988) suggests that all nouns in Taos, Kiowa, Zuni, and the Algonquian languages are inflected for number. Ojibwe, being an Algonquian language, is thus expected to grammaticize the mass/count distinction. The aim of this section is to show that this is indeed the case.

While many properties associated with number in Halkomelem Salish appear (at least at first) to indicate that it is derivational across the board, the status of number in Ojibwe is clearly inflectional. For instance, contrary to Halkomelem Salish, number in Ojibwe is obligatory (section 10.2.1), triggers agreement (section 10.2.2), is not possible inside compounds (section 10.2.3) or inside derivational morphology (section 10.2.4). Moreover, Ojibwe has pluralia tantum (section 10.2.5) and bare plurals (section 10.2.6), two properties lacking in Halkomelem Salish and which Wiltschko (2008) argues follow from her analysis.

Wiltschko (2007, 2008) proposes that number must be parameterized cross-linguistically. While number in English is inflectional (it is projected in the syntax, DP selecting a #P in the nominal domain), as shown in (5a), it is non-inflectional in Halkomelem Salish (#P is not projected), as illustrated in (5b). Instead, plural marking in that language is a modifier attached to a nominal root (i.e. it is derivational).

![Diagram](5.png)

While this captures the differences between languages such as English and Halkomelem Salish, it does not say anything about languages such as Ojibwe where number seems to be inflectional, but where certain mass nouns can nevertheless be pluralized. The parameter also appears to be too strict, since in many Indo-European languages where number is clearly inflectional, certain mass nouns can nevertheless be pluralized (see section 10.3.2).

4 The term pluralizer is the term used by Wiltschko (2008) for modificational plural markers.
Rather than viewing inflectional morphology and derivational morphology as two different systems (use of number as the head of #P in one language versus use of number as a modifier to the nominal root in another language), I will follow Booij (1993, 1995) in viewing both types of morphology as processes (the distinction between inflectional and derivational morphology is far from clear, see for example Distributed Morphology, Halle and Marantz 1993, Marantz 1997, 2001 where the distinction is blurred).

Wiltschko (this volume) proposes another parameter that caters for languages like Blackfoot, another Algonquian language, since it has pluralized mass nouns but clear inflectional uses of number otherwise. I will review this new parameter in section 10.3.

10.2.1 Plural marking is obligatory in Ojibwe

In English, plural marking is obligatory. For example, the numeral three, which forces a plural interpretation, triggers plural marking on the noun (6a). The unmarked (singular) form is not compatible with a plural interpretation as illustrated by (6b).

(6) a. the three boy-s
   b. *the three boy

Wiltschko (2008) shows that Halkomelem Salish (HS) is different: a noun following a numeral of cardinality greater than one can but need not be marked for plural. In (7) plural marking is optional: the unmarked form is compatible with a plural interpretation, a characteristic of general number (Corbett 2000). In other words, there is no meaning difference associated with the presence of the plural marker.5

(7) a. te lhíxw swíweles a'. te lhíxw swóweles [HS]
   det three boy det three boy.pl
   'the three boys' 'the three boys'
   b. qex te s-th’im b’ qex te s-th’eth’im
   many det nmz-berry many det nmz-berry.pl
   'many berries’ 'many berries’

(Wiltschko 2008:642)

5 The Halkomelem Salish plural has the following three forms, usually treated as phonologically conditioned allomorphs: reduplication, -l- infixation, and vowel change (Ablaut).

(i) a. reduplication: mèle mámele
   child children
   b. -l-infixation: q’ámi q’álemi
   girl girls
   c. Ablaut: swíweles swóweles
   boy boys

In Ojibwe, on the other hand, pluralization is obligatory just as in English, as shown by the examples in (8a,b). In Ojibwe, plural nominals receive a different kind of plural marking depending on whether they are animate (-ag) or inanimate (-an). As already mentioned in the introduction, Ojibwe is a language where gender is based on animacy.6 The unmarked noun (boy or gwiizens) is interpreted as singular and is thus incompatible with a plural interpretation. (8a’,b’) are not grammatical because plural marking is missing. The plural morpheme is necessary for a plural interpretation. This is a typical property of inflectional plural marking.

(8) a. niizh gwiizens-ag a’. *niizh gwiizens [Ojibwe]
   two boy-PL_ANIM two boy
   ‘two boys’ intended: ‘two boys’

b. niibina miin-an b’. *niibina miin
   many berries-PL_IN many berry
   ‘many berries’ intended: ‘many berries’

10.2.2 Plural marking triggers agreement

Another pervasive property of inflectional plural marking is its ability to trigger agreement. For example, a noun marked for plural in English requires that the preceding demonstrative be marked for plural as well (9a). If the demonstrative is not marked as plural, the sentence is ungrammatical as shown by (9b). Conversely, an unmarked (singular) noun is not compatible with plural marking on the demonstrative (9c); rather, an unmarked noun must be preceded by a singular demonstrative (9d).

(9) a. These boys can sing.
   b. *This boys can sing.
   c. *These boy can sing.
   d. This boy can sing.

Halkomelem Salish is again different from English. As shown by Wiltschko (2008), a noun marked for plural is compatible with a preceding plural determiner (ye in (10a)) but it does not require a plural determiner. Rather, a plural-marked noun is also compatible with the unmarked determiner te (as in (10a’)). Similarly, an

6 While Hallowell (1960), Black-Rogers (1982), Straus and Brightman (1982), and Corbett (1991:23) have argued that gender in Algonquian languages has a semantic basis (i.e. the crucial notion being power), it appears in fact that the animate/inanimate contrast has grammaticized over time. The gender system cannot be entirely semantic since some inanimate elements receive the animate gender agreement: aakim ‘snowshoe’, seeman ‘tobacco’, miskomin ‘raspberry’, etc. While miskomin ‘raspberry’ is animate, ‘strawberry’ odeminin is inanimate, which shows how unpredictable gender can be in the language. There is also considerable variation between dialects indicating that gender in Ojibwe is grammatically based. The gender shift phenomenon described later in the chapter shows that animacy has often nothing to do with power, but more to do with measure or unit readings. In addition, the idea that gender in Algonquian is tied to power should not be exaggerated (see footnote 15).
unmarked noun is compatible with plural marking on the determiner (10b) but it can also co-occur with the unmarked determiner (10b').

(10) a. t’ilém ye s-í:wí:qe. a’. t’ilém te s-í:wí:qe. [HS]
    sing DET.PL man.PL sing DET man.PL
    ‘The men are singing.’ ‘The men are singing.’

b. t’ilém ye swíyeqe. b’. t’ilém te swíyeqe.
    sing DET.PL man sing DET man
    ‘The man is singing.’ ‘The man is singing.’

(Wiltschko 2008:643)

Assuming that plural marking in Halkomelem does not instantiate a functional head and is thus not capable of triggering the syntactic operation Agree, Wiltschko (2008) predicts that there is no agreement for number established elsewhere in the grammar. For example, in Halkomelem, subject-verb agreement does not include number, only person. In the case of 3rd person, the same subject agreement marker (-es) is used for both singular and plural subjects, as illustrated in (11).

    help-TRANS-3S DET.PL man.PL DET.PL woman.PL
    ‘The men are helping the women.’

b. máy-t-es te swíyeqe ye slheláli.
    help-TRANS-3S DET man DET.PL woman.PL
    ‘The man is helping the women.’

(Wiltschko 2008:654–655)

In Ojibwe, the situation is different. Agreement is obligatory between the noun and the preceding demonstrative as shown in (12) for animates and in (13) for inanimates. (12b’) and (13b’) are ungrammatical because the noun does not carry plural marking.

(12) a. maaba gwinzens
    this.anim boy
    ‘this boy’

b. maamig gwinzens-ag b’. *maamig gwinzens
    these.anim boy-pl.anim these.anim boy
    ‘these boys’ intended: ‘these boys’

(13) a. maanda baagan
    this.in nut
    ‘this nut’

b. maamin baagan-an b’. *maamin baagan
    these.in nuts-pl.in these.in nut
    ‘these nuts’ intended: ‘these nuts’
It is noteworthy that, in Ojibwe, subjects agree differently with the verb depending on whether the subject is singular or plural. Compare (14a) with (14b). In (14a), the suffix -ig is the 3rd person singular agreement while in (14b) the suffix -oog is the 3rd person plural agreement.\(^7\)

(14) a. nene n-gii-waabm-ig.  
   man-1SG 1SG-PAST-see-3SG  
   '(A) man saw me.'

   b. nenwag n-gii-waabm-igoog.  
      men-3PL 1SG-PAST-see-3PL  
      'Men saw me.'

10.2.3 Plural marking is not possible inside compounds in Ojibwe

As is well-known, plural marking in English is prohibited inside of compounds as illustrated by the examples in (15).

(15) a. baby-sitting  a’. *babies-sitting

   b. key-ring  b’. *keys-ring

As shown by Wiltschko (2008), the situation is different in Halkomelem Salish. Although English-type compounds, where two free forms are merged, are rather rare, it is often possible for a bound root to merge with a free root thus forming a compound (Gerdts and Hinkson 1996, Gerdts 1999, 2003). Plural marking is possible on the non-head of such compounds, as shown on the basis of the following kind of head initial compounds.

(16) a. tem-qoqo: qo = water  
      time-water.pl
      'high water time'

   b. tem-weléxes wéxes = frog
      time-frog.pl
      'time of frogs’ (=‘March’)

(Wiltschko 2008:644)

Similarly, the initial root within such a compound can also be pluralized without necessarily pluralizing the referent of the complex noun. For example, in (17) the exocentric compound can refer to one chipmunk with more than one stripe.

\(^7\) Note that, as is well known, Ojibwe proclitics do not agree with subjects or objects but with the argument which is higher on a person/gender hierarchy: 2nd > 1st > Animate 3rd > Obviative (Animate) 3rd > Inanimate. Third person marking is usually not pronounced. When it is, it is suffixal rather than a proclitic.
In Ojibwe, on the other hand, plural marking is not possible inside of compounds, as shown by (18a) for animate nouns and (18b) for inanimate nouns (compounds are head last).

(18) a. aamoo-ziinzibaakwad a’. *aamoo-g-ziinzibaakwad
   bee-sugar bee-pl.anim-sugar
   ‘honey’
   b. ishkode-daaban b’. *iskode-n-daaban
   fire-car fire-pl.in-car
   ‘train’

If the nominal is pluralized, the whole compound is necessarily pluralized (19). Thus, the situation is unlike that of Halkomelem Salish, but similar to English.

(19) ishkode-daaban-an
    fire-car-pl.in
    ‘trains’

10.2.4 Plural marking is not possible inside derivational morphology

Another property often used to identify inflectional categories has to do with their distribution relative to derivational morphology. Generally, inflectional categories are impossible inside of derivational morphology, as illustrated by the English examples in (20). A root nominal is first merged with a derivational affix. Only then, an inflectional affix can be merged with the complex word as in (20a–d). The reverse in (20a’–d’) is impossible.

(20) a. dog-ish a’. *dog-s-ish
    b. mother-ese b’. *mother-s-ese
    c. brother-hood c’. *brother-s-hood
    d. tattoo-ist d’. *tattoo-s-ist

In contrast, plural marking in Halkomelem Salish is possible inside of derivational morphology. As shown in (21), plural reduplication appears between the nominalizing prefix s- and the root.

(21) a. s-p’eq’p’eq’ s-p’eq’ = white
    NMZ-white.pl
    ‘white spots on skin’
    b. s-th’ekw’th’ekw’ s-th’eth’ikw’ = ‘be sore’
    NMZ-cont.sore
    ‘sore’, ‘lots of sores’

(Galloway 1993:379, in Wiltschko 2008:645)
On the other hand, Ojibwe has a series of nominalizers -gan, -win, -w (among others), which cannot appear outside plural marking. Plural number morphology is never possible between the nominalizer (traditionally taken to be a derivational morpheme) and the root, as illustrated by the following examples.8

\[\text{(22) a. } \text{bkwenzh-gan-ag} \quad \Rightarrow \text{‘breads’} \]
\[\text{b. } \text{wazas-win-an} \quad \Rightarrow \text{‘nests’} \]

10.2.5 Ojibwe has pluralia tantum

Wiltschko (2008) argues that, since number is not a grammatical category in Halkomelem Salish, no mismatches between form and meaning can arise. The presence of a plural modifier in Halkomelem Salish must be interpreted. This contrasts with English plural marking which functions as a syntactic head and as such can display form-meaning mismatches. This is most apparent in pluralia tantum. In English, pants and scissors are plural but refer to single objects. Wiltschko (2008) reports that in Halkomelem Salish there are no cases of pluralia tantum (Galloway 1993).

On the other hand, just like English, Ojibwe has a few nouns that are inherently plural. Valentine mentions (23a–b) and a plural participial noun (23c).

\[\text{(23) a. } \text{biiwekdamaagnan} \text{ ‘wood shaving (W)} \]
\[\quad \text{[Ojibwe]} \]
\[\text{b. } \text{bootsan} \text{ ‘boots’ (from English boots, Odawa)} \]
\[c. \text{e-baasgobjigemgakin ‘spring-tooth harrows’, a farm implement. (Valentine 2001:182)} \]

10.2.6 Ojibwe has bare plurals

Since on Wiltschko’s (2008) analysis, plural marking in Halkomelem is merged as a modifier, it is predicted that in that language the plural marking does not change the category of the noun it merges with. In English, on the other hand, a pluralized noun is no longer a noun, but instead functions as a number phrase (#P). In Halkomelem Salish a pluralized noun is still a noun and does not become a #P. This structural difference suggests that in English, but not in Halkomelem, pluralized nouns have a distribution distinct from unmarked nouns. The idea is that English pluralized nouns can be used as arguments without a preceding determiner (i.e. as so-called bare plurals). This is illustrated in (24).

---

8 This may in fact not be the best test to use, because the apparent possibility for inflection to feed derivation also occurs in languages where plural marking is clearly otherwise inflectional: the case of compounds in Dutch (Booij 1993, 1995) and of compounds in Breton (Stump 1991). In Ojibwe, contemp- tives can appear before or after plural marking (Valentine 2001).
(24)  a. I saw bears.
     b. Bears saw me.

Following Déchaine and Wiltschko (2002), Wiltschko (2008) assumes that the presence of some functional structure (such as #) is sufficient to turn a nominal predicate into an argument. Consequently, the well-formedness of English bare plurals in argument position reflects the presence of functional structure (i.e. #P). Nouns without such superstructure are not licensed as arguments. In Halkomelem the plural marker does not change the syntactic category of the pluralized noun, i.e. it does not project functional structure. As a consequence, Halkomelem does not have bare plurals: all nouns in argument position must be preceded by a determiner (see Matthewson 1998 for a cross-Salish perspective). This is illustrated in (25).

(25)  a. tsel kw'ets-l-exw *(te) {sw̃iyeqe/si:wi:qe}. [Halkomelem Salish]
     1sg.s see-trans-3o det man/man.pl
     'I saw the man/the men./' 'I saw a man-men.'
     b. t'ít'elem *(te) {sháli/slhelhláli}.
     singing det woman/woman.pl
     'The woman/women is/are singing.'/ 'A woman/women is/are singing.'

In Ojibwe, bare plurals are possible in both subject and object position as attested by the examples in (26). This means that #P is projected and that #Ps can be used directly as arguments without the need of a determiner.

(26)  a. n-gii-waabmag nenwag. [Ojibwe]
     1sg-past-see-3pl.anim men-3pl.anim
     'I saw men.'
     b. nenwag n-gii-waabm-igoog.
     men-3pl.anim 1sg-past-see-3pl.anim
     'Men saw me.'

10.2.7 Intermediate summary

Table 10.1 summarizes the findings above:

<table>
<thead>
<tr>
<th></th>
<th>English</th>
<th>Halkomelem</th>
<th>Ojibwe</th>
</tr>
</thead>
<tbody>
<tr>
<td>Obligatory plural marking</td>
<td>yes</td>
<td>no</td>
<td>yes</td>
</tr>
<tr>
<td>Obligatory agreement</td>
<td>yes</td>
<td>no</td>
<td>yes</td>
</tr>
<tr>
<td>Plural inside compounds</td>
<td>no</td>
<td>yes</td>
<td>no</td>
</tr>
<tr>
<td>Plural inside derivational morphology</td>
<td>no</td>
<td>yes</td>
<td>no</td>
</tr>
<tr>
<td>Pluralia tantum</td>
<td>yes</td>
<td>no</td>
<td>yes</td>
</tr>
<tr>
<td>Bare plurals</td>
<td>yes</td>
<td>no</td>
<td>yes</td>
</tr>
</tbody>
</table>
It should be clear by now that the distribution of number in Ojibwe patterns with languages such as English rather than with languages such as Halkomelem Salish. Ojibwe is a language with inflectional number. As seen in (27), the opposition between singular and plural cuts across the paradigms of nouns, pronouns and verbs. Agreement is obligatory with all these categories.

(27)  

<table>
<thead>
<tr>
<th>SINGULAR</th>
<th>PLURAL</th>
<th>GLOSS</th>
</tr>
</thead>
</table>
| a. giin | giinwaa | 'thou (sg) ~ you (pl)'
| wiin    | wiinwaa | 'he ~ they'            |
| maanda  | nanda   | 'this (one) in proximal ~ these (ones) in proximal' |
| n-…-im | n-…-naan| 'my ~ our (exclusive)' |
| b. giwe | giwewag | 'go home (pres, 1/2/3 sg) ~ go home (pres, 3pl)' |

Despite all the differences between Halkomelem Salish and Ojibwe with regard to the status of number, there is, nevertheless, one remaining property that Ojibwe shares with Halkomelem Salish. Mass nouns can be systematically pluralized in both languages. Moreover, as the next section will show, mass nouns in both languages can appear with numerals or quantifiers otherwise uniquely paired with count nouns. We thus have a paradox: on the one hand, everything points to the view that in Ojibwe, number is a functional head rather than a modifier (number marking is obligatory, it triggers agreement, and it is not possible inside compound or derivational morphology). On the other hand, the fact that many mass nouns can be pluralized might be taken to indicate that number in Ojibwe is not an inflectional category and that a grammaticized mass/count distinction is absent in the language. The next two sections address this problem.

10.3 Ojibwe pluralized mass nouns

In section 10.2, we saw that the number system of Ojibwe shares many properties with the number system of English and that contrary to Halkomelem Salish, number marking in Ojibwe is obligatory, triggers agreement, and is not possible inside compound or derivational morphology. This indicates that number in Ojibwe is an inflectional category. In the simplest cases, pluralization in Ojibwe thus creates a series of discrete individuals. In (28a), ‘nuts’ means ‘more than one nut’ and the same goes for bagaan versus bagaanag in Ojibwe. (28b–c) are other examples of animate nominals while (29) introduces a list of inanimate nouns (IN).

(28)  

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>a. bagaan</td>
<td>‘nut’</td>
<td>~</td>
<td>bagaan-ag</td>
</tr>
<tr>
<td>b. miigwan</td>
<td>‘feather’</td>
<td>~</td>
<td>miigwan-ag</td>
</tr>
<tr>
<td>c. maanadikoshens</td>
<td>‘goat’</td>
<td>~</td>
<td>maanadikoshens-ag</td>
</tr>
</tbody>
</table>
What is noteworthy in Ojibwe is that mass nouns that cannot be pluralized in English can easily be pluralized and be associated with numerals and quantifiers otherwise uniquely used with count nouns. We turn to this in more detail.

10.3.1 Pluralized mass nouns in Ojibwe

As already mentioned in the introduction, mass nouns in Ojibwe can easily be pluralized. This is true for animate nouns and for inanimate nominals.

In English, the pluralization of these nouns is simply not possible (unless a kind interpretation is targeted, see footnote 1). Some nouns in this list are collective rather than mass nouns. This is the case, for example, of (30f).

9 The examples in (31) involve pluralized mass nouns that are inanimate. It is expected under my analysis that they should be animate. However, it may be that these examples simply exhibit the continuing loss of the singulative system. For example, in some dialects the plural of 'rice' is animate while in others it is inanimate. It appears that for some speakers, the singulative system is no longer productive or that at least there are major lexical gaps.
These nominals cannot even be pluralized to mean ‘kinds of’, for example ‘oils’ as ‘kinds of oil’; this is an observation that coheres with the quote introduced in (1) and which was confirmed by my informants. For example, speakers reject the direct translations of ‘What kind of sugars do you have?’ (for instance, upon entering a shop). It turns out that Ojibwe has a special word that it uses for kinds, call it a kindifier, namely *dnawa* or *dowa* (cf. Valentine 2001:593, my consultants used *daawa*).

In addition, the pluralized nouns in (32) resist a conventionalized unit reading. This shows that the alleged systematic pluralization of mass nouns in Ojibwe is not a kind of generalized coercion mechanism and that the concept of conventionalized unit is orthogonal to the issue of mass noun pluralization. Although mass nouns cannot normally be pluralized in English, exceptions to that generalization are possible provided that the interpretation of mass nouns is coerced to that of kinds (33a) or standard servings (33b).

(33) a. There are only three waters available (tap, still, and sparkling water).
   b. John ordered three waters (i.e. glasses, bottles etc.).

The mass/count distinction in English, although grammaticized, is therefore actually quite flexible. Mass nouns can be made count (via the Universal Sorter, e.g. (33a), or via the Universal Packer, e.g. (33b), see Bunt 1985b) and count nouns can be made mass (David Lewis’s Universal Grinder, e.g. *there was dog all over the road*, cf. Pelletier 1979).10

For Ojibwe, it is not difficult indeed to think of ‘milks’, ‘teas’, ‘waters’ or ‘oils’ as conventionalized units in the Ojibwe community (bottles of oil, cartons of milk, cups of tea). Yet, these particular mass nouns cannot be pluralized in the language.

Conversely, many pluralized mass nouns in (31) clearly do not involve conventionalized units; a fact which again indicates that in Ojibwe pluralization of mass nouns is a separate process from that of conventionalized unit creation. When ‘ice’, ‘mud’, ‘grass’, ‘air’, ‘moss’ are pluralized, there is no sense that the portioning leads to a conventionalized unit, since there is no conventionalized unit that corresponds in the Ojibwe culture to these nominals.

---

10 In fact, as pointed out by a reviewer, even in English it is not always clear that coercion is entirely flexible. In other languages, there is definite evidence that coercion is syntax-dependent. The same reviewer points out that the ‘dog-stuff’ reading is not possible in Mandarin Chinese (Cheng, Doetjes and Sybesma 2008). In that language, there is no syntactic reason to coerce since bare singulars are fine in argument position.
In sum, the grammatical distinction between mass and count terms is not neutralized in the language. What is noteworthy is that all the pluralized mass nouns in (30) and (31) receive a measure reading. Semaag means ‘wads or pieces of tobacco’, mikomiig ‘pieces of ice’, mashkosiwag ‘blades of grass’, akiin ‘bits of earth’, bkwezhganan ‘slices or pieces of bread’, wiiyaasan ‘pieces or cuts of meat’, etc. When the noun is a collective rather than a mass noun, we obtain a simple unit reading as in the case of zhoomiyag ‘coins’. Once we fully appreciate the meaning behind the pluralization of mass nouns in Ojibwe, it is easier to understand why mass nouns in Ojibwe can appear with exactly the same numerals and quantifiers that are otherwise associated with count nouns. I turn to this in the next section.

10.3.2 Mass nouns and numerals/quantifiers

In English, while count nouns can be modified by cardinal numerals, e.g. *three nuts, mass nouns cannot, e.g. three mud(s). Also, English count nouns can be modified by quantifiers such as many, few, every and each, e.g. many/few nuts, every/each nut. On the other hand, mass nouns cannot be modified by such quantifiers: *many/few mud(s), every/each mud.

In Ojibwe, mass nouns can not only be pluralized – see examples in (30) and (31) – they can, just like count nouns (34), also be modified by cardinals (35) and can be modified by the same quantifiers used for count nouns. Compare (36) with (37).

(34) a. bezhig baagan b. niizh baagan-an [Ojibwe]
    one nut two nuts.pl.in
    ‘one nut’ ‘two nuts’

(35) a. bezhig azhashki b. niizh azhashki-n
    one mud two mud.pl.in
    ‘one chunk of mud’ ‘two chunks of mud’

(36) a. gakina baagan b. gakina gwiizens
    every nut every boy
    ‘every nut’ ‘every boy’

(37) a. gakina azhashki b. gakina ziinzibaakwad
    every mud every sugar
    ‘every piece of mud’ ‘every piece of sugar’

Halkomelem Salish behaves exactly like Ojibwe in this regard. The quantifier qex (many/much) can be used with mass nouns (38) as well as with count nouns (39).

(38) a. tsel kwétslexw qex (te) qó/qoqo.
    1sg.s see-trans-3o qu det water/water.pl.
    ‘I have seen lots of water.’
\[b. \text{el stl’i kw qex (te) mélk/memlk.}\]
\[\text{1sg.poss want det qu det milk/milk.pl}\]
\[\text{‘I want lots of milk.’} \quad (\text{Wiltschko 2007})\]

\[39\]
\[\text{a. tsel kw’éts-lexw qex (te) theqá/theqtheqát.}\]
\[\text{1sg.s see-trans qu det tree.pl}\]
\[\text{‘I saw lots of trees.’}\]
\[\text{b. tsel kw’éts-lexw qex (te) sth’im/sth’eth’im.}\]
\[\text{1sg.s see-trans qu det berry}\]
\[\text{‘I saw lots of berries.’} \quad (\text{Wiltschko 2007})\]

It is also possible to combine numerals with both count and mass nouns as shown by
the examples in (40) and (41) respectively.

\[40\]
\[\text{a. tsel kw’éts-l-exw isále sth’im/sth’eth’im.} \quad [\text{Halkomelem Salish}]\]
\[\text{1sg.s see-trans-3o two berry/berry.pl}\]
\[\text{‘I have seen two berries.’}\]
\[\text{b. tsel kw’éts-l-exw isále theqát/theqtheqát.}\]
\[\text{1sg.s see-trans-3o two tree/tree.pl}\]
\[\text{‘I have seen two trees.’} \quad (\text{Wiltschko 2007})\]

\[41\]
\[\text{a. tsel kw’éts-l-exw isále siyítsem.}\]
\[\text{1sg.s see-trans-3o two sand.pl}\]
\[\text{‘I have seen two pieces of/kinds of sand.’}\]
\[\text{b. tsel kwéts-l-exw isále siyólh}\]
\[\text{1sg.s see-trans-3o two wood.pl}\]
\[\text{‘I saw two pieces of wood.’} \quad (\text{Wiltschko 2007})\]

If we focus on the interpretation that these pluralized mass nouns get in Halkomelem Salish it is easy to notice that there is an important difference in interpretation between Halkomelem and Ojibwe. In Ojibwe, plural marking leads to a reading where the mass term is divided into portions or pieces (the measure reading). In contrast, the function of the plural with mass nouns in Halkomelem Salish is simply, in most cases, to denote a large amount of mass (without a quantifier, e.g. qex). Consider the following examples.

\[42\]
\[\text{a. tsel kw’éts-l-exw te/ye th’exth’ëxet.} \quad [\text{Halkomelem Salish}]\]
\[\text{1sg.s see-trans-3o det/det.pl gravel.pl}\]
\[\text{‘I saw a lot of gravel.’}\]

\[11\] Ojibwe is not the only Native North-American language with plural mass nouns triggering a unit of measure reading. Mithun (1999:80) mentions the case of Yup’ik: ‘Yup’ik mass nouns such as \textit{iqquq} “oil” and \textit{meq} “water” often appear with number suffixes to indicate units of substance, such as “containers of”: \textit{iqqu}-\textit{k} “oil-dual” = “two sealpokes of oil”; \textit{mer}-\textit{et} “water=plural” = “buckets of water.” Yup’ik appears to have all the properties associated with number in Ojibwe: obligatory, disjunctive, transparent, word final, yet it allows the pluralization of mass nouns.
b. tsel kw’êts-l-exw te/ye syiqyiq.
1sg.s see-trans-3o det/det.pl snow.pl
‘I’ve seen a lot of snow.’

c. tsel kw’êts-l-exw te/ye spepíw.
1sg.s see-trans-3o det/det.pl ice.pl
‘I’ve seen a lot of ice.’

d. tsel kw’êts-l-exw te/ye shweláthetel.
1sg.s see-trans-3o det/det.pl fog.pl
‘I’ve seen a lot of fog.’

(Wiltschko 2008:669)

There are a couple of exceptions, see the examples in (41). In these cases, the translation by Wiltschko (2007) is ‘pieces of’. However, this reading is shown to be possible only when numerals are used, not when determiners are used. In most cases, it seems that, in order to divide mass terms, Halkomelem Salish resorts to partitive quantifier phrases of the kind available in English. The element i’axwlí which literally means ‘small’ (43a) is used as such a phrase. An example is introduced in (43b).

(43) a. i’axwíl siyólh [Halkomelem Salish]
   piece of wood

b. tsel kw’êts-l-exw (te) i’axwíl siyólh.
   1sg.s see-trans-3o det small wood
   i) ‘I saw a piece of wood.’
   ii) ‘I saw a little bit of wood.’

On the other hand, Ojibwe does not appear to have measure phrases corresponding to ‘a piece of’, ‘a bit of’ or ‘a portion of’.12 As Greenberg (1972:16) has made explicit, ‘there are a considerable number of Amerind languages […] which do not have measure constructions [including Ojibwe, Greenberg 1974b]. Numerals occur directly both with nouns designating mass as well as countable objects.’ This is an important point and a crucial fact of the language which explains why mass nouns can be pluralized in Ojibwe with an atomizing reading. We turn to this in more detail.

10.3.3 The interpretation(s) of pluralized mass nouns

The main point that I would like to make is that the different uses of pluralized mass nouns as we find in Ojibwe on the one hand, and in Halkomelem Salish, on the other, as well as the different interpretations tied to these different uses, are independent from the parameter proposed by Wiltschko (2007, 2008). The pluralization of mass

12 Measure words are not completely absent: the word for ‘pound’ is diibaabiishkojian.
nouns is often possible in languages with number as inflectional. This is clear in more familiar languages of the Indo-European stock: French (44), Hebrew (45), Persian (46), and also Biblical English (47).

(44) a. La fonte des neiges/les neiges éternelles [French]
   the melting of the snows/the snows eternal
   ‘The melting of the snow’ / ‘The eternal snow(s)’
   b. Les eaux du Nil ont débordé de leur lit.
   the waters of the Nile have spilled from their bed
   ‘The waters of the Nile have spilled over from their bed.’
   c. Des viandes avariées gisaient sur la table.
   some meats bad lay-IMP on the table
   ‘Bad meats were lying on the table.’

(45) Iarah harbe šeleg/šlagim. [Hebrew]
   fell.3SG.PAST a lot snow.sg.masc/ snow.pl.masc
   ‘A lot of snow fell/has fallen.’

(46) åba-e/åb-å-ye daryâ bålâ umad-an. [Persian]
   water-ez/ water-pl-ez sea high came
   ‘The sea level rose.’
   (Sharifan and Lotfi 2003:231)

(47) And he said unto them, Go. And they came out, and went into the swine:
   and behold, the whole herd rushed down the steep into the sea, and perished
   in the waters. (Matthew 8:32)

A special interpretation is attached to all the examples where the mass noun is plural. These plurals appear neither dividing nor counting. Rather, they appear to denote abundance: see, for example, Corbett (2000) about the plural of abundance. The plural of abundance is one of several plurals (plural of modesty, exaggerative plural, hyperbolic plural, approximative plural, anti-associative plural, etc.) that has nothing to do with counting individuals. These plurals often can take special forms. For example, in Banyun, a language of the West Atlantic branch of Niger-Kordofanian, spoken in Senegal and Guinea Bissau, the plural of abundance (sometimes called the greater plural) is used when the exact number is impossible to pinpoint or when it is irrelevant, and it surfaces as a different form from that of the dividing plural. In that language, nouns typically have singular and plural, distinguished by prefixes, but the special plural is used when the number cannot be counted or the speaker feels it unnecessary (Corbett 2000, Sauvageot 1967).

In the French example illustrated in (44a), although an aspectual reading is clearly available as well (something like the annual melting of the snow is implied), la fonte

13 Thanks to Keren Tonciulescu for the Hebrew data.
des neiges ‘the melting of the snow’ means abundance of snow. In Persian, it appears that the more spreading, pouring, or gathering a context indicates, the more the use of the plural form (Sharifan and Lotfi 2003, Hamedani 2011). The judgements are subtle, and often abundance is not necessary to obtain the reading: it is sufficient for the pluralized mass noun to denote a vague notion of number (number is either difficult to pinpoint or unimportant to the speaker).

This appears to be the kind of reading we obtain with Halkomelem Salish pluralized mass nouns. If this is correct, number in Halkomelem might thus be inflectional after all, the pluralization of mass nouns arising from the use of inflectional number inherently in addition to its more traditional contextual use. The distinction between inherent and contextual inflectional morphology was introduced by Booij (1993, 1995) who shows that in certain languages (he concentrates on Dutch), in addition to the traditional use of number where a verb agrees with a nominal subject, it is possible to use number in a derivational fashion. Rather than relegating number to derivational processes, however, Booij (1993, 1995) argues that it must be kept part of inflectional morphology but with the added twist that number can also be used inherently. Inflectional morphology and derivational morphology are not two different kinds of morphology, but two different processes. When used inherently it might be used as a modifier as in Wiltschko (2008) or introduced in the nominal structure via n (cf. Lecarme 2002, Kihm 2005, Acquaviva 2008).

In other words, there is no need to claim that number in Halkomelem Salish is derivational across the board. Number in that language, is never, after all, impossible, but simply optional. In addition, many of the tests used by Wiltschko (2007, 2008) would equally work in languages with inflectional number (see footnote 8).

To summarize: in this section, we saw that the abundance or non-counting reading yielded by pluralized mass nouns is available in languages where number is clearly inflectional and that this use of the plural is thus completely independent from whether the language has a grammaticized count/mass distinction. In Ojibwe, we saw that pluralized mass nouns yield a reading that is either a simple unit or a unit of measure. We must now explain how this unit reading is made possible in the language. I turn to this in the next section.

10.4 Basis for a solution

In the preceding sections, we established that number in Ojibwe has clear inflectional properties and that the reading obtained when mass nouns are pluralized is one that corresponds to individuation or measuring. On the other hand, Halkomelem Salish pluralized mass nouns were shown to yield a different kind of interpretation, i.e. one that appears to be related to abundance or vagueness of number. In this section, I provide a solution to the Ojibwe puzzle: why are pluralized mass
nouns possible in the language and why is a unit interpretation the reading obtained?

The main idea introduced in this section is that the unit of measure reading yielded by pluralized mass nouns in Ojibwe stems from the use of gender shift in the singular to express singularization. The same mechanism is used to create individuals out of collectives. Pluralized mass nouns are possible because gender shift has applied in the singular prior to pluralization. This will be shown to be the reflex of a singulative system. Before I give details about the singulative, I want to discuss briefly the status of gender in Ojibwe.

Ojibwe, like other Algonquian languages, has a productive gender system based on animacy. Some nouns are animates while others are inanimate. The animate/inanimate contrast is grammaticized (like the masculine/feminine gender system of, say, Romance languages): for example, some types of berries are inanimate while others are animate, some body parts are inanimate while others are animate, etc. (see footnote 6).

At this point, it is worth reviewing Wiltschko’s (this volume) claim that animacy in Blackfoot (and thus perhaps in other Algonquian languages) is not gender but nominal aspect. I will argue that in Ojibwe animacy is definitely gender. Wiltschko (this volume) claims on the basis of Blackfoot that animacy is not gender, but nominal aspect. Following Rijkhoff (1991), Wiltschko makes an analogy between the verbal and the nominal domain in terms of boundedness. Bounded events have an inherent endpoint/culmination, beyond which the same event cannot continue (e.g. walk to the store) while unbounded events have no inherent endpoint/culmination—the same event can continue over an indefinite period of time (e.g. play cards). For the nominal domain, the idea is that mass nouns are unbounded (they are not collections of individuals) while count nouns are bounded (they form collections of individuals).14

Wiltschko’s (this volume) idea is that English has a clear grammaticized mass/count distinction, therefore nominal aspect is instantiated by a bounded versus unbounded distinction. On the other hand, Blackfoot does not, according to Wiltschko, have a grammaticized mass/count distinction (Wiltschko cannot find examples of mass terms that cannot be pluralized in that language). This means that the language lacks a nominal aspect instantiated via a bounded versus unbounded contrast. Instead, in Blackfoot nominal aspect is instantiated via animacy. Compare (48a) with (48b).

\[(48)\]
\[\text{a. } [\text{DP } \text{D}^0 [\text{NumP } \text{Num}^0 [\text{AspP } \text{Asp}^0 [nP n^0 [\sqrt{N}]]]]] \quad \text{[English]} \]
\[\downarrow \]
\[\text{[+/− bounded]} \]

14 This is an idea that has been popular in French for French mass nouns (see Anscombe 1986).
In order to support her claim, Wiltschko uses the fact that in Blackfoot, verbs have different forms depending on whether the participants are animate or inanimate. She argues that verbal Aktionsart is thus not based on [+/- bounded]/telicity, but instead that in Blackfoot Aktionsart is based on [+/- animate]. Wiltschko also argues specifically that animacy in Blackfoot is not like gender. In German, she shows that nominalizers are specified for gender but that in Blackfoot they are not: German die Grammat-ik ‘the grammar’ (feminine) versus der Lehr-er ‘the teacher’ (masculine); the Blackfoot nominalizer a'tsis is used for both animate nouns (saa’kssoya-a’tsis ‘poison ivy’) and inanimate nouns (isoohkama-a’tsis ‘container’).

In Ojibwe, the count/mass distinction is clearly grammaticized, as I have shown in the present chapter. For example, not all mass nouns can be pluralized: liquids resist pluralization. Also, while it is true that verbs in Ojibwe end in either animate or inanimate forms—for example, the verb ‘be’ has two forms (see also the examples in (49) below): one for animate nouns (-wi), and another for inanimate nouns (-wan), cf. Valentine (2001), this is in addition to a [+/- bounded] distinction. As pointed out by Valentine (2001), verbs such as those in (49) can be interpreted not only as ‘be’ but also as ‘become’ where a change of state is clearly asserted.

(49) a. ninii-wi. [Ojibwe]
   man-be
   ‘He is a man.’ or ‘He has become a man.’

b. oodenaa-wan.
   town-be
   ‘It is a town.’ or ‘It has become a town.’

More generally, Ojibwe verbal expressions corresponding to ‘walk to the store’ are bounded while ‘play cards’ are not. In sum, telicity is, as in English, a concept very much anchored in the grammar of Ojibwe.

Finally, while it is true that Ojibwe nominalizers are indifferent as to whether the noun they merge with is animate or inanimate (e.g. the nominalizer -gan can appear with either), it is not clear how this is relevant to distinguishing languages with a grammaticized count/mass distinction from languages without one. In conclusion, there is no reason in Ojibwe to treat animacy as (nominal) aspect instead of the more traditional notion of gender. Everything points to the view that animacy is nothing other than gender in Ojibwe.
Let me now focus on gender shift. While it is well-known in the Algonquian literature that it is possible in Algonquian languages to shift gender from inanimate to animate as a mark of elevated status, especially in story-telling (a case of semantic shift, Valentine 2001:118, Black-Rogers 1982, or perspectival marking, Muehlbauer 2008), it is less known that it is also possible to shift gender from inanimate to animate as a way to divide what would otherwise be a mass or collective noun. In this case, we get either an individual reading as in (50) or a unit of measure reading as in (51). The examples are from Fox (Mesquakie), cf. Goddard (2002). -i is the inanimate marking on nouns while -a is the animate marking. Goddard (2002) does not correlate these facts with the notion of the singulative, but I believe this is exactly what is at work here.

(50) a. zhooniyaahi ‘silver, money’ IN ~ zhooniyaaha ‘a coin, a bill’ ANIM [Fox]
   b. miichipehi ‘game’ IN ~ miichipeha ‘a game animal’ ANIM

(51) a. owiiyaasi ‘meat, flesh’ IN ~ owiiyaasa ‘a piece/cut of meat’ ANIM
   b. owiinenwi ‘fat (generic)’ IN ~ owiinenwa ‘a piece of fat’ ANIM
   c. anakehkwi ‘bark’ IN ~ anakehkwa ‘a piece of bark’ ANIM

(Goddard 2002:213)

Once singularized, nouns such as zhooniyaaha (ANIM) ‘silver, money’ can be pluralized, giving us zhooniyaaha-ki (ANIM) ‘coins, bills, money’. The same goes for miichipehi (IN) ‘game (collective)’ → miichipeha (ANIM) → miichipeha-ki (ANIM) ‘game animals’.

It is evidently difficult to notice the gender shift that is the spell out of the singulative in Ojibwe. This is because Ojibwe nouns have lost their final -i for inanimates and -a for animates for most singulars. There are, however, visible effects of this process in modern Ojibwe as I am about to show.

First, let me point out that there is evidence, as argued by Piggott (2007), that number is present in the derivation of every Ojibwe noun. Each of the singular forms in (52a–c) ends in a vowel that is demonstrably not part of the exponent of the root morpheme. The root allomorphy in (52c) [miʃ [mis] results from a palatalization process (s → ŋ) that only applies in a derived environment (Kaye and Piggott 1973). This means that there is a singular suffix -i that is attached to inanimate nouns and an animate counterpart -a. Since singular number and gender are fused morphologically, it is reasonable to assume that gender is also present in the derivation of every Ojibwe noun.

15 This is a separate phenomenon and in my view not the realization of the singulative. Inanimates can become animates if they become important in the story or if they acquire power. However, as pointed out by Goddard (2002), the use of gender shift in narratives to express power should not be exaggerated. For example, in Fox inanimates seem to be freely assigned the powers of speech, comprehension, and thought without shifting gender (Goddard 2002:208).
In words where no final -i or -a surfaces it is assumed that the vowel has been truncated. Interestingly, however, as argued by Piggott (2007), the vowel can only be truncated if the word meets minimality requirements. If the word is too small, e.g. it is bisyllabic, the vowel cannot be truncated. The process is, therefore, systematic and predictable. The word nika ‘goose’ in (54a) is clearly singular and animate and its plural version nikag ‘geese’ is visibly plural and animate.

For most singular nouns, however, especially those that end with a consonant, it is impossible to tell whether they are animate or inanimate from the endings, creating a situation where the singulative is not morphologically visible. In addition, because some words have lost their final -i together with final consonants, inanimate nouns may end in an -a which is the mark for animates rather than inanimates (and vice-versa).

With this in mind, take the case of the collective noun zhoonya ‘money’ in (53a). It is listed in dictionaries as an inanimate noun (although it ends in -a), but interestingly in the plural it is listed as animate. The same goes for mitig in (53b). It is listed in dictionaries as inanimate when it is interpreted as ‘wood/forest’, but as animate when it means ‘tree’. Although there is no morphological evidence, this is a clear residual effect of the gender shift correlated with the singulative use (in Fox, zhooniyaha ‘silver, money’ ends in -i the mark of the singular inanimate while zhooniyaha ‘a coin, a bill’ ends in -a the mark of the singular animate).

In order to show that nouns such as zhoonya ‘money’ and mitig ‘tree’ are animate on their individual reading we can introduce them in a sentence. Since Ojibwe has two different kinds or verb stems, animate transitive (TA) versus inanimate transitive (TI), the kind of verb stem selected will depend on the gender of the noun. Moreover, animate nouns are marked obviative. Ojibwe distinguishes between two third persons in a sentence or a narrative by means of a mechanism called obviation. In the sentence ‘John saw Fred’, for example, there are two third persons, John and Fred. When a sentence contains two third persons in this kind of grammatical relationship, one of them is seen as the main one and is called proximate (as if it were somehow closer to the interest of the speaker) and the other one is seen as secondary and is called obviative. In Ojibwe, the obviative is marked only on animate nouns. In (54a) makwa ‘bear’ is animate; it takes the obviative form -n
and the verb stem is TA. In (54b) *jiiman ‘boat’ is inanimate; it takes no obviative form and the verb stem is TI.

(54) a. John o-gii-waabam-aa-n makwa-n. [Ojibwe]
   John 3SG.s-PAST-see.anim-3SG.o-OBV bear-OBV
   ‘John saw a bear.’

   John 3SG-PAST-see.in-3SG.o boat boat-OBV
   ‘John saw a boat.’

When we turn to mass and collective terms, we see that in (55a), mnoomin ‘rice’ is obviative and the verb stem is TA and that the same goes for zhoonya ‘money’ in (55a).

(55) a. John o-gii-waabam-aa-n mnoomin-an agidisag. [Ojibwe]
   John 3SG.s-PAST-see.anim-OBV rice-OBV on.floor
   ‘John saw a grain of rice on the floor.’

b. John o-gii-waabam-aa-n zhoonya-an agidisag.
   John 3SG.s-PAST-see.anim-OBV coin-OBV on.floor
   ‘John saw a coin on the floor.’

We can now see why, in Ojibwe, mass terms can systematically be pluralized. It is not because the language lacks a grammaticized count/mass distinction, but because in these cases, a gender shift process has operated on the singular to give a plural with a gender that is different from that of the original mass/collective noun. When the input to the atomization process is a collective noun we obtain a unit reading and when the input is a mass noun we obtain a unit of measure reading (pieces of x, portions of x, blades of x—depending on the noun).

The singulative is an important component of the grammars of Breton, Welsh, Arabic, Dagaare, etc. It can turn a mass or collective term into a noun denoting an individual or unit of measure reading. In these languages, the change from mass/collective to individual/unit of reading surfaces with a shift of gender from masculine to feminine. My point in this chapter is that in Ojibwe this is exactly what happens: gender shift (from inanimate to animate) is a reflex of the singulative. That Algonquian grammars might embed a singulative system has, to the best of my knowledge, never been entertained.

As an individualizer, the Breton singulative can target collective nouns as shown in (56) (examples from Stump 2005: 62). These are semantically plural, but morpho-syntactically singular. The output nominal is individuated. Breton is like Ojibwe: number marking is obligatory, triggers agreement, etc. It is clearly inflectional. It is in fact always the case, it seems, that languages with a singulative system, have grammars with number as an inflectional category. There do not appear to be
languages with only derivational number (contrary to what is predicted by Wiltschko’s 2007, 2008 parameter).

(56) 

a. buzhug ‘worms’ \(\sim\) buzhug-enn ‘a worm’

b. kraon ‘walnuts’ \(\sim\) kraon-enn ‘a walnut’

c. per ‘pears’ \(\sim\) per-enn ‘a pear’

d. logod ‘mice’ \(\sim\) logod-enn ‘a mouse’

The resulting singulative form can in turn be pluralized as shown in (57).

(57) 

a. buzhug-enn ‘a worm’ \(\sim\) buzhug-enn-où ‘worms’

b. kraon-enn ‘a walnut’ \(\sim\) kraon-enn-où ‘walnuts’

c. per-enn ‘a pear’ \(\sim\) per-enn-où ‘pears’

d. logod-enn ‘a mouse’ \(\sim\) logod-enn-où ‘mice’

The process is productive and appears to be syntactic in the sense that it occurs in exactly in the same syntactic contexts as the choice between an ordinary singular noun and its plural counterpart. For example, Stump (2005) shows that the syntactic contexts that determine the choice of the ordinary singular noun \(potr\) ‘boy’ (lenited form \(botr\)) and its plural counterpart \(potred\) ‘boys’ likewise determine the choice between the singulative noun \(sivienn\) ‘strawberry’ (lenited form \(zivienn\)) and its collective counterpart \(sivi\) ‘strawberries’. This is shown in Table 10.2 (from Stump 2005: 63).

When mass nouns are the target of the singulative operation we get a unit of measure reading.

(58) 

a. douar ‘earth, ground’ \(\sim\) douar-enn ‘plot, terrier’

b. geot ‘grass’ \(\sim\) geot-enn ‘blade of grass’


\[\text{Table 10.2. Stump (2005:63)}\]

<table>
<thead>
<tr>
<th>(potr) ‘boy’</th>
<th>(sivi) ‘strawberries’</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Singular contexts</strong></td>
<td><strong>Singular: potr</strong></td>
</tr>
<tr>
<td></td>
<td><strong>ur potr bennak</strong></td>
</tr>
<tr>
<td></td>
<td>‘a certain boy’</td>
</tr>
<tr>
<td></td>
<td>meur a botr</td>
</tr>
<tr>
<td></td>
<td>‘many a boy’</td>
</tr>
<tr>
<td></td>
<td>Plural: potred</td>
</tr>
<tr>
<td></td>
<td>un nebeud potred</td>
</tr>
<tr>
<td></td>
<td>‘some boys’</td>
</tr>
<tr>
<td></td>
<td>kalz potred</td>
</tr>
<tr>
<td></td>
<td>‘a lot of boys’</td>
</tr>
</tbody>
</table>
The function of the singulative thus consists not only in turning abstract object types into identifiable objects, but also picking discrete entities out of undifferentiated mass. As pointed out by Acquaviva (2008: 245), the precise sense of the unit of measure reading varies with the word. This state of affairs is widespread in languages that allow the singulative and is exactly what we find in Ojibwe (see above).

In Mathieu (forthcoming), I show that Borer’s (2005) system can accommodate the singulative on the assumption that in addition to classifiers and the plural, division can be performed by the singulative. Borer (2005) proposes that all nouns in all languages are in need of being portioned out before they can interact with the count system. In Chinese, this is achieved by (count) classifiers for count nouns and (mass) classifiers for mass nouns (Cheng and Sybesma 1999). In English, plural marking takes the role of classifiers for count nouns and measure phrases are used in lieu of mass classifiers (Sanches and Slobin 1973, Doetjes 1996, 1997). For singular individual nouns in English, e.g. one cat, the numeral one functions both as a divider and a counter. In Hungarian, Tagalog and Turkish this is generalized to all numerals in which case it is possible for the noun to be free of plural marking (Borer 2005). The singulative is another way of dividing undivided stuff.

The mass/count distinction in Ojibwe

The incorporation of the singulative into Borer’s (2005) system of division is perfectly natural. While the singulative has been less studied in recent years in the context of research on number, its importance in the grammar of many languages, including, as I have shown in this chapter, Ojibwe, shows that the singulative deserves a central place in the theory of division.
10.5 Conclusion

The aim of this chapter was to show that the count/mass distinction is grammaticized in Ojibwe, that its grammar imbeds a singulative system and that the singulative is performed via gender shift (from inanimate to animate). This explains why mass nouns can be pluralized. While at first this property might be seen as a sign that the language has no grammaticized mass/count distinction, this is only an illusion.
Counting and classifiers*

LISA LAI-SHEN CHENG

11.1 Introduction: Three puzzles

We normally think of classifier-languages such as Chinese to be different from languages like English, simply because these languages use classifiers even for count nouns. Furthermore, that bare count nouns in Chinese can appear in argument positions suggest that even bare count nouns behave like mass nouns. These facts in Chinese have led to various claims concerning the interpretation of bare nouns as well as the nature of the classifiers. Below, I first put forth three puzzles in relation to the interpretation of bare count nouns, as well as the asymmetries concerning classifiers. With these puzzles as background, I re-examine classifiers in Mandarin and Cantonese. Let us consider first the Universal Grinder puzzle.

11.1.1 The Universal Grinder puzzle

Most nouns in English, for example, can be either mass or count, depending on the context.

(1)  a. There is steak all over the floor.
     b. Kim put apple in the salad.

(Pelletier 1979, Pelletier and Schubert 1989)

Both steak and apple in (1) have the ‘ground’ reading (i.e. being interpreted as a mass noun). This can be attributed to the so-called ‘universal grinder’. A ‘universal grinder’ (the term due to David Lewis) takes an object corresponding to any (apparent) count noun (e.g. man), and puts the object in one end of the grinder, and asks what is on the floor (e.g. There is man all over the floor). Interestingly, for a language like (Mandarin) Chinese, though many have claimed that it only has mass

* I would like to thank Susan Rothstein, Rint Sybesma, Jenny Doetjes, Denis Delfitto, Pino Longobardi, and Roberto Zamparelli for their help and stimulating discussions. I thank Joanna Sio for discussing her Cantonese judgements with me, and Yiya Chen and Chinhui Lin for discussing their Mandarin judgements with me.
nouns (see Borer 2005, Chierchia 1998a,b among others), corresponding grinder examples (as in (2a)) do not lead to the same reading as we saw in (1).1

(2) a. qiáng-shang dōu shì gōu. [Mandarin]  
   wall-top all cop dog  
   ‘There are dogs all over the wall.’  
   not: ‘There is dog all over the wall.’

b. qiáng-shang dōu shì gōu-ròu.  
   wall-top all cop dog-flesh/meat  
   ‘There is dog(meat) all over the wall.’

c. di-shang dōu shì shuǐ.  
   floor-top all cop water  
   ‘There is water all over the floor.’

Crucially, (2a) has the so-called ‘wall-paper’ reading (see Cheng, Doetjes and Sybesma 2008). The noun gōu ‘dog’ does not have the reading that would have resulted from a universal grinder. Instead, the sentence yields the picture that the wall has a wall-paper with (little) dogs on it.

Given the contrast between English and Chinese, we have the following ‘universal grinder puzzle’: if all nouns in Chinese have a mass denotation, how come gōu ‘dog’ in (2a) cannot have a mass interpretation? Note that in the Chinese case, it is as if we cannot even appeal to the universal grinder. Instead, the sentence yields the picture that the wall has a wall-paper with (little) dogs on it.

11.1.2 Classifier reduplication puzzle 1: Mandarin vs. Cantonese

Cheng (2009a) shows that there is a systematic difference between Cantonese and Mandarin in the reduplication of classifiers.2 (3a,b) show that Mandarin classifiers cannot be reduplicated, in contrast with Cantonese (4a,b); de and ge in the glosses (here and elsewhere in this article) stand for the particular type of classifier in the examples.

(3) a. *Ge-ge rén dōu yōu zījì de lixiāng. [Mandarin]  
   cl-cl person all have self de ideal  
   ‘Everyone has his own ideal.’

---

1 I will gloss dōu as ‘all’ in this chapter. But see Cheng 2009a, which treats dōu as a maximality operator.

2 Typical sortal classifiers are glossed as cl, while mensural classifiers are glossed as cl with its meaning put as a superscript. The numbers in the Cantonese examples indicate tones.
b. *Ge-ge chúshī dōu zuò yī-dào cài.
   chef all make one-dish
   'Every chef makes a dish.'
   (data adapted from Yang 2004)

(4) a. Go³-go³ jan⁴ dou¹ jau⁵ zì⁶ ge³ le³ so⁵ se⁴⁵. [Cantonese]
   person all have self ideal
   'Everyone has his own ideal.'

b. Go³-go³ cyu² dou¹ zuo⁶-zo jat¹-dip⁶ sung³.
   chef all make-perf one-dish
   'Every chef makes a dish.'

Given that both Mandarin and Cantonese are classifier languages, what is the difference between the classifiers in the two languages which can lead to such a difference in reduplication?

11.1.3 Classifier reduplication puzzle 2: Cantonese

In (4), we see that Cantonese classifiers can be reduplicated. However, there is some restriction on reduplication. As shown in (5) and (6), the reduplication of measure phrases is restricted. The question arises as to why (5b, c) are ungrammatical, while (6) is grammatical.

(5) a. bong⁶-bong⁶ yuk⁶ dou¹ hou² san¹ sin¹. [Cantonese]
   pound-pound meat all very fresh
   'Every pound of meat is fresh.'

b. *ma⁵-ma⁵ bou³ dou¹ hou² leng³.
   yard-yard cloth all very pretty
   'Every yard of cloth is very pretty.'

c. *cek³-cek³ dei⁶ dou¹ hou² gon¹ zeng⁶.
   foot-foot floor all very clean
   'Every foot of floor is very clean.'

(6) cek³-cek³ bou⁶ dou¹ jat¹ ye⁶-ye⁶ gam³ fut³. [Cantonese]
   cloth all same such wide
   'Every foot of cloth is all the same width.'

In this chapter, I address these puzzles. I will first review the arguments that the count/mass distinction is still found in Chinese, though not at the nominal level; but rather, at the classifier level. In sections 11.3 and 11.4, I examine classifiers in Mandarin and Cantonese further. After discussing the differences between Mandarin and Cantonese, I turn to further examine a set of classifiers which Cheng and Sybesma (1998) call 'massifiers' (see below); in particular, I re-examine massifiers in
relation to the *de*-test, the adjective test, as well as their ability to reduplicate. I show that massifiers do not behave uniformly with respect to the tests. In section 11.5, I discuss *di*\(^1\) in Cantonese, which can be considered to be a plural classifier as well as the implication this has for our understanding of bare nouns and the nature of classifiers.\(^3\)

### 11.2 Count/mass at the classifier level

Cheng and Sybesma (1998, 1999, 2005) argue that classifiers are not all the same. Following Tai and Wang (1990) and Croft (1994) among others, they made a distinction between classifiers that create a unit of measure and the ones that name the unit in which the entities denoted by the noun come naturally. They call the first type massifiers and the second type count-classifiers.

They employed two tests to distinguish these two types of classifiers: their co-occurrence with *de* (which is typically considered to be a modification maker (see also Cheng and Sybesma 2009 for a difference analysis of *de*)), as well as their co-occurrence with adjectives such as *small* and *big*. They suggest that the difference among classifiers reflects a count-mass distinction. Below we review each of these tests in turn.

#### 11.2.1 Classifier + *de* (Mandarin)

The classifiers that are associated with ‘count’-nouns (i.e. count-classifiers) cannot be followed by *de* (such as (7)), whereas container classifiers or measure classifiers (i.e. massifiers) can (as in (8)) (see also Chao 1968, Paris 1981, Zhu 1982, Tang 1990).\(^4\)

---

\(^3\) In this chapter, I do not discuss the phenomenon in which a classifier appears following the noun. See Zhang (this volume) for a detailed discussion. I do not think that when the ‘classifier’ appears post-nominally, it is a functional category on a par with a prenominal classifier. As Zhang (this volume) notes, its distribution is similar to a bare noun. Furthermore, typical count-classifiers cannot all appear postverbally (e.g. *yī-zhāng zhuōzǐ* [one-CL-table] vs. *zhúō-zhāng yī-zhī gōu* [one-CL-dog] vs. *gōu-zhī*) and in some cases, the meaning changes when the putative classifier appears postverbally (e.g. *yī-fēng xīn* ‘one-CL-letter’ vs. *xīn-fēng* ‘envelope’), the latter fact leads one to think of this as something along the lines of compounding, which can yield non-compositional readings.

\(^4\) Li (2008) briefly discusses a couple of counter-examples as in (i) and (ii):

(i) *shí duò běn de shū*  
`ten more CL volume de book`  
‘approximately 10 books’

(ii) *liàng bāi duò fēng de xīn*  
`two hundred more CL de letter`  
‘approximately 200 letters’

Note that all these examples involve a number marked with *duò* ‘more’, and it gives an approximate number. Hsieh (2008) also discusses such examples. Her conclusion is that in cases when sortal classifiers can appear with *de*, either the quantity is approximate, or there is contrastive focus on the classifier.
Aside from the fact that they can appear with de, massifiers differ from count-classifiers in that they can occur with both count (a set) and mass nouns (as in (8b)), while count-classifiers can only appear with count nouns. (9) provides the Cantonese counterparts, showing that massifiers can appear with ge3, the Cantonese counterpart of de.

As discussed in Cheng and Sybesma (1998), when de appears with a container/measure phrase, it provides a quantity reading. Sān-bēi shuǐ [three-CLcup water] can have a reading in which the three cups are present (such as ‘she is holding three cups of water in her hands’), as well as a quantity reading, as in ‘you need to put three cups of water in the soup’. However, when de is present, as in sān-bēi de shuǐ [three-CLcup de water], the non-quantity reading is not available.
11.2.2 Adjective + classifier

Massifiers and count-classifiers further differ in their co-occurrence possibility with *dà* 'big' and *xiǎo* 'small'.\(^5\)\(^6\) (10)–(11) show that massifiers can co-occur with these adjectives while count-classifiers cannot.

\[(10)\]
\[
a. \ yi \ dà \ zhāng \ zhī \text{ [Mandarin]}
\]
\[
\text{one big } cl^{\text{sheet}} \text{ paper}
\]
\[
\text{‘one large sheet of paper’}
\]
\[
b. \ nà \ yi \ xiǎo \ xiàng \ shū
\]
\[
\text{that one small } cl^{\text{box}} \text{ book}
\]
\[
\text{‘that one small box of books’}
\]

\[(11)\]
\[
a. \ *yi \ dà \ zhī \ gǒu \text{ [Mandarin]}
\]
\[
\text{one big } cl \text{ dog}
\]
\[
b. \ *yi \ dà \ wèi \ lǎoshī
\]
\[
\text{one big } cl \text{ teacher}
\]

Cheng and Sybesma (1998, 1999) argue that the count-mass distinction is not reflected at the lexical level in Chinese languages, but at the classifier level.

Casting this in an approach as Rothstein (2010), we can say that the mapping between natural atomicity and semantic atomicity is at the classifier level in Chinese languages, while in English it is at the lexical level (through a lexical operation) (natural atomicity being inherently individuable while semantic atomicity is atomicity relative to a context \(k\)). This provides an answer to the second question in section 11.1.1, namely, if Chinese has a count/mass distinction, why are classifiers used even for count nouns? There is a generalized use of classifiers in Chinese languages because the mapping between natural atomicity and semantic atomicity is not at the lexical level; instead, the mapping takes place at the classifier level. It is therefore very important to understand what classifiers really are.

---

\(^5\) Zhu (1982:52) indicated that some ‘count-classifiers’ can also have such adjectives preceding them. The examples that he gave involve nouns such as *zhī* ‘paper’, *shílòu* ‘stone’, *fěizào* ‘soap’, and *bīng* ‘ice’. However, all of these nouns are arguably ‘mass’ nouns. It should also be noted that in verifying whether or not an adjective can be added, one should also vary the numeral, and not limit it to one only. With the numeral one, in some cases, it does not have a numeral reading, as the numeral one is in some cases comparable to the English article *a*. And in other cases, it provides a reading similar to ‘a whole’.

\(^6\) Li (2008) puts forth some counter-examples to Cheng and Sybesma (1998, 1999). However, it should be noted that the examples all have a special reading. Consider (i):

\[(i)\]
\[
\text{wǔ-mǎo qián mǎi-le yi dà ge mángguó}
\]
\[
\text{50-cent money buy-perf one big } cl \text{ mango}
\]
\[
\text{‘Such a sizable mango only costs 50 cents.’}
\]

Again, in such cases, when the numeral is changed to a higher one than one, the sentence becomes degraded.
11.3 Chinese classifiers

Consider first the following list of classifiers.

(12) a. yī-běn shū
    one-cl book
    ‘a book’

    b. yī-jiàn jiānjù
    one-cl piece furniture
    ‘a piece of furniture’

    c. yī-kuài dàngāo
    one-cl slice cake
    ‘a slice of cake’

    d. yī-bēi shuǐ
    one-cl cup water
    ‘a cup of water’

    e. yī-shēng shuǐ
    one-cl litre water
    ‘a litre of water’

On the surface, the list in (12) appears to provide the same information: the nouns in Mandarin are preceded by a classifier when we have a numeral. And if there is a difference, we expect a difference along the lines discussed above concerning massifiers versus count-classifiers. However, I show below that among the so-called massifiers, there is a difference in their behavior with respect to the two tests mentioned above. Before we discuss this, we need to first turn to consider the difference between Cantonese and Mandarin, since it provides us with a window to the nature of classifiers.

11.3.1 Cantonese vs. Mandarin

Cheng and Sybesma (1999) note that there is a difference between Cantonese and Mandarin classifier-noun combinations. In particular, whereas Mandarin bare nouns can be used to denote definiteness, Cantonese bare nouns cannot. Instead, to express definiteness in Cantonese, classifier-noun combinations (without a numeral) are used (rather than bare nouns). This is illustrated in (13) and (14). (14a) illustrates that the bare noun in Mandarin is interpreted as definite in a bounded event. However, in the same environment in Cantonese, a classifier-noun sequence has to be used (13a). In (14b), a bare noun in Mandarin appears in the subject
position, and it is obligatorily interpreted as definite. In contrast, (13b) shows that in the same environment in Cantonese, the classifier has to be present.\(^7\)

(13) a. Wu\(^4\)fei\(^1\) jam\(^2\)-jyun\(^4\) *(wun\(^2\)) tong\(^1\) la. [Cantonese]
   Wufei drink-finish CL\(^{\text{bowl}}\) soup SFP
   ‘Wufei finished drinking the soup.’

   b. *(Zek\(^3\)) gau\(^2\) gam\(^1\)jat\(^6\) dak\(^6\)bit\(^6\) teng\(^1\)waa\(^6\).
   CL dog today special obedient
   ‘The dog is specially obedient today.’

(14) a. Hufei hê-wán-le tâng. [Mandarin]
   Hufei drink-finish-LE soup
   ‘Hufei finished the soup.’

   b. *(zhī) gōu jíntiān tèbié tīnghuà.
   CL dog today very obedient
   ‘The dog/dogs was/were very obedient today.’

The difference with respect to classifier-nouns and bare nouns between Cantonese and Mandarin is not restricted to expressing definiteness. Sybesma (2008) notes that we see the same effect with specific indefinites. (15) and (16) provide examples where specific indefinites are facilitated. We see again that in Cantonese, in such cases, the classifier must be present, and in these cases in Mandarin, the classifier is optional.

(15) a. lo\(^2\) *(tiu\(^4\)) sing\(^2\) bong\(^2\)-sat\(^6\) leung\(^5\) zek\(^3\) geok\(^3\). [Cantonese]
   take CL rope bind-tight two CL legs
   ‘Bind both legs tight with a rope.’

   b. yòng (gēn) shéngzi bǎ iâng zhī tuì bāng-shàng. [Mandarin]
   use CL rope ba two CL leg bind-up
   ‘Bind both legs up with a rope.’

(16) a. zik\(^1\)-hak\(^1\) pai\(^3\) *(go\(^3\)) din\(^6\)-gung\(^1\) lei\(^4\). [Cantonese]
   immediately send CL electrician come
   ‘Send an electrician over immediately’.

   b. mǎshâng pài (ge) diângōng lái. [Mandarin]
   immediately send CL electrician come
   ‘Send an electrician over immediately’.

In Cheng and Sybesma (1999), it is stipulated that Mandarin classifiers cannot be used without a numeral. The presence of numerals is associated with (non-specific) indefinites. Sybesma (2008) starts to address the difference between Mandarin and

\(^7\) In both Mandarin and Cantonese, we see classifier-noun combinations in the case of indefinites (in object positions of an unbounded event for instance). In these cases, it is possible that there is a null numeral one present (see Cheng and Sybesma 1999).
Cantonese classifiers. He first discusses the difference in the two languages in terms of
the use of classifiers. In particular, he points to Erbaugh (2002), which shows that
(a) more nouns appear without a classifier in Mandarin than in Cantonese; (b) the
number of specific classifiers (as opposed to the general one) used by Cantonese
speakers is far higher than the number used by Mandarin speakers; and (c) in
Mandarin, the general classifier ge is used much more often than its counterpart
go3 in Cantonese.

Sybesma (2008) then shows that while hundreds and hundreds of the most
common nouns in Mandarin feature the suffix -zi, the counterpart of this element
in Cantonese is lacking with the same nouns. The Dàoxù Xiândài Hânyǔ cídiān
('Reverse dictionary of Modern Chinese [Mandarin]') lists close to one thousand
nouns suffixed by -zi. Note that -zi has two functions; one is a nominalizer, as in the
example lóng-zi [deaf-zi] ‘deaf person’; the other function is more difficult to
pinpoint. Consider the examples in (17).

(17) a. hái-zi ‘child’ [Mandarin]
    b. zhuō-zi ‘table’
    c. fáng-zi ‘house’

Hái ‘child’, zhuō ‘table’, and fáng ‘house’ in (17) are the so-called bound mor-
phemes, which typically do not appear alone. If they do not appear with -zi, they
have to appear with another morpheme, as shown in (18).

(18) a. xiăo-hái ‘child’ [Mandarin]
    small-child
    b. shū-zhuō ‘desk’
    book-table
    c. shū-fáng ‘study-room’
    book-house

Going through more than 600 such nouns, Sybesma notes that these nouns are
virtually all count nouns (only 17 may be questionable). This link between -zi and
count nouns has already been observed by Rygaloff (1973). The corresponding nouns
in Cantonese are bare, without zi, as shown in (19).

(19) a. saǐ³ Lou⁶ ‘child’ [Cantonese]
    b. tòi² ‘table’
    c. uk¹ ‘house’

In other words, large numbers of count nouns in Mandarin are marked with -zi,
while in Cantonese, no such marking is present. Sybesma concludes from this that
count nouns in Mandarin come out of the lexicon marked as count (e.g. by -zi),
while count nouns in Cantonese are not marked as such in the lexicon.
11.3.2 Two classifier nodes

Based on the difference between Cantonese and Mandarin in terms of the use of -\textit{zi} and the use of the general classifier, Sybesma (2008) proposes that -\textit{zi} is a unit marker; that is, it marks a count-noun as count (from the lexicon). That is, count-nouns in Mandarin come out of the lexicon with a unit marker, marking the nouns as count. It therefore follows that a count-classifier in Mandarin is not a unit marker. It simply allows a numeral to attach to a noun (see Doetjes 1997, as well as Cheng and Sybesma 1999). Let’s call such classifiers Cl-N(\textit{umeral}) here. This can thus provide an explanation as to why classifiers in Mandarin have to appear with a numeral—without a numeral, classifiers are simply not needed in Mandarin. On the other hand, Cantonese count-nouns do not come out of the lexicon marked as count. The count-classifier in Cantonese thus plays the role of a unit marker. Let’s call such classifiers Cl-U(\textit{unit}). Its appearance is thus not restricted to the presence of a numeral.

In other words, there are two types of count-classifiers, namely, Cl-u and Cl-n. Cl-u is a unit-marker, and Cl-n bridges the numeral and the noun. One possible way to implement this is to have two classifier projections, with the count-classifier marking units as the lower classifier projection, as indicated in (20) (see Cheng and Sybesma 2009).

\begin{center}
(20)
\begin{tikzpicture}
  \node (numep) {NumeP}
  \node[below of=numep] (numeral) {Numeral}
  \node[below of=numer] (clpn) {ClP\textsubscript{N}}
  \node[below of=clpn] (cl0-n) {Cl\textsuperscript{0}-N}
  \node[below of=cl0-n] (cl0-u) {Cl\textsuperscript{0}-U}
  \node[below of=cl0-u] (np) {NP}
  \node[below of=clpn] (clpu) {ClP\textsubscript{U}}
  \draw (numep) -- (numeral)
  \draw (numeral) -- (clpn)
  \draw (clpn) -- (cln)
  \draw (cln) -- (clu)
  \draw (clu) -- (np)
  \end{tikzpicture}
\end{center}

I hypothesize here that count-classifiers in Cantonese start out from the lower Cl-u position, as they are unit-markers and move to the higher Cl-n position, in order to bridge the numeral and the noun phrase. On the other hand, Mandarin count-classifiers start out in Cl-n, being selected by the Numeral.

Now we can turn to the Classifier reduplication puzzle 1. To recapitulate, Cantonese classifiers can easily reduplicate while the ones in Mandarin cannot. If Cantonese classifiers are unit markers, the null hypothesis is that only unit markers can reduplicate, yielding a distributive reading. The reduplication of Cl-u thus yields universal quantification over individual units; in contrast, Mandarin count-classifiers are not unit markers, and thus cannot be reduplicated (see also Cheng 2009a).
11.4 Do all classifiers individuate or divide?

In Cheng and Sybesma (1999), it is stated that ‘[l]ike D, the count-classifier may be said to have a singularizing function: the count-classifier identifies singular units; it picks out one instance of what is denoted by N’ (p. 517). In other words, there is a simple divide between count-classifiers and massifiers: count-classifiers spellout the unit denoted by the noun while massifiers create the unit for counting/measuring. The question that arises is whether such a simple divide holds. In particular, given the difference that we have seen above between Cantonese and Mandarin, it is essential that we examine the classifiers in a little more detail. I will argue that not all massifier classifiers are individuators/dividers, and even for the dividers, they do not always divide.

I have reviewed in section 11.2 the two tests used in Cheng and Sybesma (1999) for distinguishing count-classifiers from massifiers (i.e. the de-test and the adjective-test). These tests will be used below to further flesh out the distinctions among the classifiers.

11.4.1 Count-classifiers vs. containers

If a count-classifier is a Cl-u (as in the case of Cantonese), or a Cl-n, it is by definition not an individuator or a divider. A typical count-classifier simply spells out the unit that comes with the count noun. Note that these are the classifiers which cannot be followed by de or preceded by the adjectives small and big, as shown in (21) (as well as the examples in (8), (9) and (11)).

\[(21)\]
\[
a. \text{ sān (\text{*xiǎo}) zhī (\text{*de}) gǒu} \quad \text{[Mandarin]}
\text{three small Cl de dog}
\text{‘three dogs’}
\]
\[
b. \text{ saam\textsuperscript{1} (\text{*sai\textsuperscript{3}}) zek\textsuperscript{3} (\text{*ge\textsuperscript{3}}) gau\textsuperscript{2}} \quad \text{[Cantonese]}
\text{three small Cl ge dog}
\text{‘three dogs’}
\]

Contrast a count-classifier with a container-phrase. When container-phrases are used as classifiers (thus as massifiers), they can be followed by de or preceded by the adjectives small and big, or both ((8b) is repeated here as (22a)). Thus, the contrast between a count-classifier and a container-classifier is the prototypical difference between count-classifiers and massifiers stated in Cheng and Sybesma (1999).

\[(22)\]
\[
a. \text{ liǎng xiāng (de) shū} \quad \text{[Mandarin]}
\text{two Cl\textsuperscript{box} de book}
\text{‘two boxes of books’}
\]

\[
\]

---

8 I use numerals higher than one just to ensure that we are really using numerals rather than readings such as whole (see footnote 4).
b. sān dà xiàng shū
   three big box book
   ‘three big boxes of books’

(23) wǔ dà běi de jīǔ [Mandarin]
   five big cup de wine
   ‘five big cups of wine’

Note that the container-classifiers have two functions: they individuate and measure (see also the discussion above concerning the interpretation of the noun phrase with de present). We also see this in English:

(24) a. Add two cups of wine to the soup. (only measure function)
    b. Put two bottles of wine on the table. (only individuating function)

When de is present, only the measure reading is present. It is not possible to order a glass of wine in a restaurant by using (25b) ((25a) must be used instead).

(25) a. yī-běi jīǔ [Mandarin]
    one-cup wine
    b. yī-běi de jīǔ
    one-cup de wine

Note that the adjectives small and big in these cases (i.e. with measure phrase, as in (22) and (23)) modify the container phrase (that is, big boxes and big cups), and not the noun phrase itself (i.e. big books and big wine).

11.4.2 Classifiers associated with furniture nouns

Consider now a class of nouns which are in between simple count nouns and mass nouns, namely, furniture nouns, which contain naturally atomic elements (such as tables and chairs), but they are not semantically atomic, in the sense that we do not use them as count nouns. These nouns are similar to the count nouns in (21), since they are also nouns with naturally atomic elements. Consider now the classifiers that are used with furniture nouns in Cantonese and Mandarin, as in (26).

(26) a. sān-jìàn jiājù [Mandarin]
    three-piece furniture
    ‘three pieces of furniture’

    b. saam^1-gin^6 ga^1 si^1 [Cantonese]
    three-piece furniture
    ‘three pieces of furniture’

Classifiers such as jiàn/gin^6 ‘piece’ respect natural atomicity. They denote sets of semantic atoms, which are countable (such as three pieces of furniture). These classifiers do not individuate, and they also do not create a unit for counting in
the same way that *cup creates the unit for counting for wine. Consider now the examples in (27) which show their co-occurrence with de and adjectives.9

(27) a. sān dà jiàn jiājù [Mandarin]
    three big CLpiece furniture
    ‘three big pieces of furniture’

b. *sān jiàn de jiājù
    three CLpiece DE furniture

(27a,b) show that the de-test and the adjective test diverge. The classifiers which are used for furniture-nouns can be modified by small and big, though they cannot be followed by de. In other words, classifiers associated with furniture nouns differ from typical count-classifiers, which cannot be modified by big or small. However, these classifiers are not compatible with a quantity measure.

11.4.3 Nouns without natural atomicity

Turning now to nouns without natural atomicity (i.e. without inherent individuability). Aside from using container classifiers, we can use other types of massifiers, as shown in (28).

(28) sān-kuài dàngāo [Mandarin]
    three-CLslice cake
    ‘three slices of cake’

*Kuài or slice is a good case of individuating/dividing classifiers, which impose atomic structure on matter. They don’t ‘spell-out’ a unit which comes with the nouns; instead, they create a unit for counting. These individuating classifiers have the property that the individuation they involve is relatively stable over time. Once cake has been divided up into slices, the slices hold until their structure is disturbed.

As with the classifiers for furniture-nouns, we also see a splitting between the two tests that we used earlier for distinguishing count-classifiers and massifiers. Again, de cannot be used but adjectives can be used. In other words, these classifiers have more affinity with count-classifiers in that they do not provide quantity measure.

(29) a. *sān kuài de dàngāo [Mandarin]
    three CLslice DE cake

b. sān dà kuài dàngāo
    three big CLslice cake
    ‘three big slices of cake’

9 Note that for some reason, it is better to use *dà ‘big’ in case of *jiàn ‘piece’ rather than *xiǎo ‘small’. This has nothing to do with the size of furniture, since toy, which is a ‘furniture’-noun has the same result.
Aside from this type of ‘divider’-classifier, we can also use measure phrases for nouns without natural atomicity. Consider first the examples in (30) and (31).

(30) a. sān-shēng (de) shuǐ
three-CL litre DE water
‘a litre of water’

b. sān-bàng (de) ròu
three-CL pound DE meat
‘three pounds of meat’

(31) a. *sān dà shěng shuǐ
three big CL litre water
b. *sān dà bàng ròu
three big CL pound meat

These examples show a reverse pattern from the furniture-classifiers and the ‘divider’-classifiers. That is, they cannot be modified by the adjectives big and small. This shows that measure phrases cannot be modified while containers and other massifiers can. This is probably because of the fact that measures such as litre and pound are not gradable. Note further that measure phrases, like container phrases can be used with count nouns, as shown in (32). And their ability to appear with de does not alter.

(32) a. liǎng gōngjīn (de) shuǐguǒ
two CL kilo DE fruit
‘two kilos of fruit’

b. liǎng bàng (de) píngguǒ
two CL pound DE apple
‘two pounds of apples’

These measure phrases measure overall quantity not by presupposing individual parts and counting them, but by using a unit of measure which creates ‘virtual individuals’ which can be counted, but which have no individual identity. To see this more clearly, consider the English sentence in (33).

(33) a. I bought two litres of milk.

b. I bought two bottles of milk.

(33a) tells us nothing about the units of milk that you buy, in contrast with (33b) where the container phrase is used. The individual litres have no identity; they only provide us with the overall quantity.

10 A reviewer points out that for English non-gradable nouns, it is possible to use adjectives to get to an intensive reading, such as He waited one long hour (for his bride to arrive), and that in French, un petit kilo ‘a small kilo’ can have an interpretation ‘just under one kilo’. These interpretations are however not possible in Mandarin/Cantonese.
Measure phrases raise the following questions. First, are these ‘virtual’ individuals created by measure phrases really individuals or not? Second, are all measure phrases the same? To answer these questions, we need to consider data with reduplicated classifiers which can bring out more contrasts.

Consider first the ‘divider’-classifier in Cantonese, which can be reduplicated.\textsuperscript{11}

\begin{center}
\begin{tabular}{l}
(34) faai\textsuperscript{3}-faai\textsuperscript{3} dan\textsuperscript{6} gou\textsuperscript{1} dou\textsuperscript{1} hou\textsuperscript{2} daai\textsuperscript{6}. [Cantonese] \\
\text{CL}\text{slice}\text{-CL}\text{slice} cake all very big \\
\text{‘Every slice of cake is very big.’}
\end{tabular}
\end{center}

We have suggested above that only Cl-\textit{u} can be reduplicated because they are unit-markers. The fact that (34) is grammatical indicates that the ‘divider’-classifier can be syntactically located in Cl-\textit{u}, though semantically they still divide. In other words, both dividers and unit markers can be mapped onto Cl-\textit{u}, therefore allowing reduplication (which yields universal quantification over individual units).

Consider now the sentences in (35), which show that not only is it the case that measure phrases are not all the same when it comes to reduplication ((35a) vs. (35b), (5a) vs. (5b)), but the same measure phrase can be sometimes reduplicated and sometimes not ((35b) vs. (35c)).

\begin{center}
\begin{tabular}{l}
(35) a. ?*ma\textsuperscript{5}-ma\textsuperscript{5} bou\textsuperscript{3} dou\textsuperscript{1} hou\textsuperscript{2} leng\textsuperscript{3}. [Cantonese] \\
\text{CL}\text{yard}\text{-CL}\text{yard} cloth all very pretty \\
\text{‘Every yard of cloth is very pretty.’}
\end{tabular}
\end{center}

\begin{center}
\begin{tabular}{l}
(35) b. cek\textsuperscript{3}-cek\textsuperscript{3} bou\textsuperscript{3} dou\textsuperscript{1} jat\textsuperscript{1}jeong\textsuperscript{5} gam\textsuperscript{2} fut\textsuperscript{3}. \\
\text{CL}\text{foot}\text{-CL}\text{foot} cloth all same such wide \\
\text{‘Every foot of cloth is all the same width.’}
\end{tabular}
\end{center}

\begin{center}
\begin{tabular}{l}
(35) c. ?*cek\textsuperscript{3}-cek\textsuperscript{3} dei\textsuperscript{6} dou\textsuperscript{1} hou\textsuperscript{2} gon\textsuperscript{1}zeng\textsuperscript{6}. \\
\text{CL}\text{foot}\text{-CL}\text{foot} floor all very clean \\
\text{‘Every foot of floor is very clean.’}
\end{tabular}
\end{center}

We have seen that ‘divider’-classifiers can be reduplicated (perhaps because they also appear in Cl-\textit{u}); the data in (35) seem to suggest that sometimes measure phrases appear in Cl-\textit{u}, sometimes not, which is not a very desirable conclusion.

Note that to interpret ‘every yard of cloth’ or ‘every foot of floor’ in the context of (35a) and (35c), it is enough to create ‘virtual’ individuals. That is, we do not need to have actual separated, or individuated units like \textit{a slice of cake}). In particular, in (35a) and (35c), the reading of ‘every yard of cloth’, or ‘every foot of floor’ equals ‘the whole cloth’, and ‘the whole floor’. No individuation is actually needed. On the other hand, in (35b), we need to compare ‘every foot of cloth’ in terms of its width. Thus, ‘a foot of cloth’ has to be separated from other feet of cloth. In order words,\textsuperscript{11}

\textsuperscript{11} It should be noted that in such cases, it is still not possible to reduplicate the classifier in Mandarin.
reduplication yields actual individuations, which must be compatible with the predicate. In the case of (35a) and (35c), individuation is actually not necessary and thus cannot be done.

In contrast, in (5a), in order to measure meat to get a pound of meat, we need to isolate a certain amount (i.e. a unit) (a pound of meat has to be weighed separately from the rest of the meat). If this reasoning is on the right track, it implies that measure phrases can be individuators/dividers. And when they are dividers, they behave as other dividers which are not measure phrases in being able to reduplicate.

The distinction between dividers and non-dividers can be further supported by the behavior of container-classifiers. Consider the following contrast.

(36) a. bui\textsuperscript{1}-bui\textsuperscript{1} seoi\textsuperscript{2} dou\textsuperscript{1} hou\textsuperscript{6} mun\textsuperscript{5}. [Cantonese]  
\textsc{CL}\textsubscript{cup}-\textsc{CL}\textsubscript{cup} water all very full  
'Every cup of water is very full.'

b. *bui\textsuperscript{1}-bui\textsuperscript{1} ge\textsuperscript{3} seoi\textsuperscript{2} dou\textsuperscript{1} hou\textsuperscript{6} mun\textsuperscript{5}  
\textsc{CL}\textsuperscript{cup}-\textsc{CL}\textsuperscript{cup} ge water all very full

As mentioned above, container classifiers can individuate or measure. In (36a), in order to compare every cup of water, water has to be individuated (i.e. a cup of water has to be separated from other cups of water). Thus, the reduplication is licit. On the other hand, when container-classifiers appear with ge\textsuperscript{3}, it is necessarily of the measure function (and thus not generated in Cl-v). In this case, it is not possible to reduplicate.

As for the de/ge\textsuperscript{3} test and the adjective-test, if the presence of de/ge\textsuperscript{3} provides a quantity measure, then measure phrases that yield quantity naturally allow it. But adjectives such as big or small are unlikely to be good with measure phrases since some measures just cannot be modified (e.g. a small kilo?; a big pound?). Table 11.1 provides a summary table for these two tests.

<table>
<thead>
<tr>
<th>Table 11.1. Summary</th>
</tr>
</thead>
<tbody>
<tr>
<td>Count-Cl Massifiers</td>
</tr>
<tr>
<td>Containers Furniture</td>
</tr>
<tr>
<td>divider non-divider measure</td>
</tr>
<tr>
<td>de  *  běn/bun\textsuperscript{2} bēi/bui\textsuperscript{1} jiàn/gin\textsuperscript{6} kuái/faai\textsuperscript{3} bàng/bong\textsuperscript{6} V</td>
</tr>
<tr>
<td>Adj *  V  V  V  V  *</td>
</tr>
</tbody>
</table>

In sum, we see differences among the massifiers. It is clear that the semantics of the classifiers matter when it comes to the different tests. Whether they can reduplicate or not depends on whether they are interpreted as a divider/unit-marker or not.
11.5 Plural classifiers

One of the controversial questions in Mandarin and Cantonese is whether  xiē in Mandarin and di¹ in Cantonese are ‘plural’ classifiers. This question is completely unexpected if one considers all nouns in Chinese as mass nouns. Here, I argue that di¹ in Cantonese is a better candidate for a plural classifier than xiē in Mandarin.

Consider first the data in (37) and (38). With xiē and di¹, we get a plurality interpretation, instead of singularities. These examples also show that in their distribution, xiē and di¹ are similar to other classifiers in that they appear after the numeral one and can follow a demonstrative. It should be noted that they can also appear with mass nouns (38); I will come back to this point below.

(37) a. yī xiē shū [Mandarin]
   one xiē book
   ‘a few/some books’

b. nèi xiē shū
   that xiē book
   ‘those books’

c. jat¹ di¹ syu¹ [Cantonese]
   one di book
   ‘a few/some books’

d. go² di¹ syu¹
   that di book
   ‘those books’

(38) a. yī xiē shuǐ [Mandarin]
   one xiē water
   ‘some water’

b. jat¹ di¹ seoi² [Cantonese]
   one di water
   ‘some water’

Iljic (1994) puts forth some objections concerning analyzing xiē in Mandarin as a plural classifier (see also Yang 2005). First, in Mandarin, xiē can appear with the general classifier ge, as in (39a); so it cannot also be a classifier. This objection in Iljic (1994) does not apply to Cantonese however. Di¹ in Cantonese cannot appear with the general classifier go³ (or any other classifier), as in (39b).

(39) a. yī xiē ge rén [Mandarin]
   one xiē cl person
   ‘some people’
Second, *xiē and *di1 cannot appear with numerals other than one for counting (counting strictly requires ‘non-plural’ classifiers/measure phrases, (40a,b)).

At first sight, this seems to be very problematic for treating *xiē and *di1 as a (plural)-classifier. However, this may be related to the possibility of neutralizing number opposition in the presence of numerical modification. In many languages such as Breton and Hungarian, numerals cannot combine with plural nouns (see Acquaviva 2008).

*Di1 in Cantonese further differs from *xiē in Mandarin in a couple of respects. First, though in both languages, *xiē/ *di1 can appear without the numeral one to express indefinite plural (as in the Cantonese example (41a)), in Cantonese, *di1 behaves like regular classifiers in that *di1-N can express definiteness, as in (41b).

(41) a. keoi⁵ soeng² maaï⁵ *di¹ syu¹. 
   [Cantonese]
   he want buy *di book
   ‘He bought some books.’

b. ngo⁵ deî⁶ maaï⁵-zo² *di¹ syu¹ la³.
   [Cantonese]
   we buy-perf *di book sfp
   ‘We bought the books already.’

In (41b), the books have to be known already and previously mentioned; this holds for both go³ *hok⁶saang¹ and *di¹ *hok⁶saang¹ in (42a,b), the former contains the general (singular)-classifier.

(42) a. go³ *hok⁶saang¹ hou² cung¹ ming⁴.
   [Cantonese]
   cl student very intelligent
   ‘The student is very intelligent.’

b. *di¹ *hok⁶saang¹ hou² cung¹ ming⁴.
   di student very intelligent
   ‘The students are very intelligent.’

Second, *di1-N combinations behave like other classifier-N combinations in allowing bare modifiers/possessors (i.e. without the modification marker ge⁳) to precede them (while this is not possible in Mandarin), as in (43)–(44).
Lastly, as discussed in Arsenijevic and Sio (2008) and Cheng and Sybesma (2009), classifiers can license N-ellipsis, as in (45a). In (45b), we see that \textit{di1} can also license N-ellipsis.12

\begin{enumerate}
\item \[\text{nei}^5 \text{bun}^2 \text{syu}^1 \text{bei}^2 \text{ngo}^5 \text{bun}^2 \_ \text{hou}^2 \text{ta}^2.\] [Cantonese]
\item \[\text{nei}^5 \text{di}^1 \text{syu}^1 \text{bei}^2 \text{ngo}^5 \text{di}^1 \_ \text{hou}^2 \text{ta}^2.\] [Cantonese]
\end{enumerate}

These facts together suggest that \textit{di1} in Cantonese is a classifier. When typical count-(singular)-classifiers combine with a count noun, it yields a singularity. When \textit{di1} combines with a count noun, it yields a plurality of objects. We have seen in (38) that \textit{di1} can combine with a mass noun. In such cases, it yields an amount reading.

11.5.1 \textit{The Universal Grinder and number}

Let us now turn back to the Universal Grinder puzzle. Consider again the sentence in (46) (repeated from (2)).

\begin{enumerate}
\item \[\text{qiáng-shàng dòu shì gǒu.}\] [Mandarin]
\end{enumerate}

\[\text{wall-top all cop dog}
\] ‘There are dogs all over the wall.’

not: ‘There is dog all over the wall.’

12 It should be noted that there is a difference between \textit{yì-xiè} and \textit{xìè} in Mandarin in this respect. With \textit{yì-xiè}, it appears to be able to license ellipsis, while \textit{xìè} by itself cannot. Typical classifiers can license ellipsis even when the numeral is not present.
Cheng, Doetjes, and Sybesma (2008) argue that a mass interpretation of count nouns in languages such as English is a ‘last resort’ or ‘coerced’ interpretation. In particular, count nouns in English have to be grammatically (morphosyntactically) marked as singular by a or plural -s, as in (47a); a bare noun is not licit. In the absence of such marking, morphosyntactic coercion may take place (depending on the right context), leading to a grinding interpretation (compare (47b) and (47c); and (48a) and (48b)).

    b. There are dogs all over the wall.
    c. There is dog all over the wall.

(48) a. There is a turkey in the fridge.
    b. There is turkey in the fridge.

In contrast to languages like English, bare nouns in Chinese (Mandarin or Cantonese) are unmarked for number; therefore, Chinese will be immune to morphosyntactic coercion (as is shown in (46)). (49a,b) further support this claim. Hènduō ‘a lot’ in (49a) is similar to a lot in English in that it combines both with mass nouns (hènduō bǐngqǐlín ‘a lot of ice cream’) and with count nouns (hènduō píngguǒ ‘a lot of apples’). (49a) shows that when hènduō appears with a count noun such as píngguǒ, we only get a count-reading, and no grinding takes place.

(49) a. wōmen zuòtiān chī-le hènduō píngguǒ/bǐngqǐlín. [Mandarin]
    we yesterday eat-perf many/much apple/ice cream
    ‘We ate many apples/much ice cream yesterday.’ (NOT: much apple)
    b. pánzi-lǐ yǒu píngguǒ/bǐngqǐlín.
    plate-inside have apple/ice cream
    ‘There are/is apples/*apple/ice cream on the plate.’

The lack of a mass reading is further shown in (49b). In this case, the context easily facilitates a mass reading of píngguǒ ‘apple’, but the mass reading is not available. This supports the availability of morphosyntactic coercion in the absence of morphosyntactic marking on number. Since Chinese does not mark number in the morphosyntax, the absence of count syntax will not trigger coercion.13

11.6 Conclusion

In the beginning of this chapter, three puzzles were put forth concerning the interpretation of bare nouns and the reduplication of classifiers. With respect to

13 Cheng, Doetjes, and Sybesma (2008) discuss a couple of other factors which sometimes lead to a coerced reading.
the Universal Grinder puzzle, as proposed in Cheng, Doetjes, and Sybesma (2008),
the count-mass coercion (i.e. grinder interpretation) requires a morphosyntactic
trigger. Since bare count nouns in Chinese are not marked morphosyntactically,
such coercion is not triggered by morphosyntax.

Concerning the reduplication puzzles, I suggest that only Cl-u’s (i.e. classifiers
that also play the role of a unit marker) can reduplicate, and this distinguishes
Cantonese from Mandarin, since the latter does not have Cl-u’s. On the other hand,
we see that measure phrases in Cantonese can sometimes reduplicate and sometimes
not. I suggest that reduplication yields actual individuation, which has to be com-
patible with the predicate.

We have seen that the divide between count-classifiers and massifiers is too
simple, though count-classifiers still stand apart when it comes to the de-test and
the adjective test, since they can go with neither de or the adjective dà 'big' or xiǎo
'small'. Massifiers do not behave uniformly when it comes to these two tests. From
the discussion above, we can conclude that de appears with quantity measures
(including container), while the adjectives dà 'big' and xiǎo 'small' appear with
almost all massifiers regardless of whether they divide or not. Their inability to
appear with measure phrases rests upon the nongradability of measure phrases.

Lastly, di in Cantonese appears to function as a classifier, including its ability to
license N-ellipsis. If di is indeed a classifier, then it further supports the claim made
in Cheng and Sybesma (1999) that classifiers express number. More work still needs
to be done to investigate the semantics of di, in particular, its combination with
mass nouns.
Countability and numeral classifiers in Mandarin Chinese

NIINA NING ZHANG

12.1 Introduction

The word zhi in the Chinese example in (1a) is called a numeral classifier (I will call it CL henceforth). A CL occurs with a numeral and an NP, such as, respectively, san ‘three’ and bi ‘pen’ in (1a). In (1b), di is also a CL.

(1) a. Yaoyao kanjian-le san zhi bi.
    Yaoyao see-perf three CL pen
    ‘Yaoyao saw three pens.’

   b. Yaoyao kanjian-le san di you.
    Yaoyao see-perf three CL oil
    ‘Yaoyao saw three drops of oil.’

Some languages have CLs and some do not. Some languages have the counterpart of the CL in (1b), but not that in (1a). From the English translations of the two examples in (1) we can see that English has the word drop to correlate with the CL di in (1b), but does not have a counterpart to the CL zhi in (1a). CLs like zhi are called individual CLs in Chao (1968), and CLs like di are called individuating CLs in this chapter. Languages that have both types of CLs, such as Chinese, are called CL languages.

Why do classifiers (CLs) exist in CL languages such as Mandarin Chinese? It has been widely assumed that the obligatory occurrence of a CL with a numeral and a noun in CL languages is related to the contrast between count and mass nominals. The goal of this chapter is to show that this traditional assumption is not fine-grained enough to cover the systematic contrasts of nominals in either Mandarin Chinese or other languages. Instead, I argue that two syntagmatic properties of nominals are syntactically significant: the ability of a noun to combine with a numeral directly, and the ability of a noun to be modified by a size or shape modifier.
The two newly recognized properties or features can be attested in the co-occurrence restrictions of quantifiers and CLs. It is the interactions between these two features, rather than the alleged binary contrast between count and mass, that explain various syntactic contrasts of countability, cross-linguistically. I argue that although the positive value of the first feature alone is enough to define the count status of a nominal, it is the combination of the negative values of both features that defines the mass status of a nominal. This chapter shows that the popular assertion that all nouns in Chinese are mass nouns is not accurate. Instead, all nouns in Chinese are non-count nouns, but they are further divided into mass and non-mass ones. The chapter also clarifies the distinctive functions of the CLs of CL languages.

In addition to this introduction and the summary section at the end, this chapter is composed of five main parts. Section 12.2 proposes my new approach to the count-mass contrast, based on the two features. Section 12.3 studies the two features in Chinese nouns and section 12.4 investigates the feature numerability of unit words. Section 12.5 further argues that the count and non-count contrast is syntactic, and section 12.6 shows the problems of certain current syntactic analyses of CLs.

### 12.2 Decomposing countability

#### 12.2.1 Identifying two new features syntagmatically

Since de Saussure (1916), two kinds of relationship between linguistic elements are recognized: paradigmatic and syntagmatic. A paradigmatic relationship is established by a substitution test. For instance, the three words of, by, and for establish a paradigmatic relation in the string government {of/by/for} the people, since each of them can substitute another. They may occur in the same syntactic position. A syntagmatic relationship is defined by the compatibility of co-occurring elements in the same construction, e.g. the relationship between the and people in the string the people. Paradigmatic and syntagmatic relationships have been metaphorically viewed as vertical and horizontal ones, respectively.

Many formal features such as tense and aspect of verbal expressions, gender and person of nominal expressions are defined paradigmatically. Selection features are typically syntagmatic features. For instance, the transitive verb drink c-selects a nominal, because it needs to occur with a nominal; and it s-selects a liquid-denoting nominal, because it needs to combine with this type of nominal.

Different kinds of syntagmatic relations exhibit different properties. In selection, the occurrence of the selected element is obligatory. But there are other syntagmatic relations that do not exhibit this kind of obligatoriness. For instance, gradability of adjectival expressions is defined by the possibility to occur with a degree word. In (2a), the adjective nice is gradable since it may occur with the degree word quite. (2b) tells us that the adjective next is not gradable, since it may not occur with quite.
(2) a. the quite nice book   b. the (*quite) next book

Another example of a non-obligatory co-occurrence relation is seen in the feature of agentivity. Agentivity of a verbal expression is defined by the possibility to be modified by an agent-oriented adverb. For instance, the VP shouted in (3a) is agentive since it may occur with the agent-oriented adverb deliberately, and the VP arrived in (3b) is not agentive, since it may not occur with deliberately.


In both the gradability and agentivity cases, a feature of an element is defined simply in the way that it allows X. Allowing does not mean requiring. Therefore the presence of X is not obligatory.

With this background in mind, I now introduce two features which are also defined syntagmatically, in order to analyze the count-mass contrast of nominals.

Some nominals may combine with a cardinal numeral directly, and some may not. In (4a), for instance, the noun apple combines with the numeral one directly. In (5a), however, the noun oil may not combine with the numeral.

(4) a. one apple   b. five apples   c. zero apples
d. 0.5 apples  e. 1.0 apples  f. five nouns

(5) a. (*one) oil  b. (*one) furniture

The contrast can also be seen in predication (adapted from Chierchia 2010: 104):

(6) a. The boys are at least thirty.  b. *The gold is at least thirty.
c. The gold is at least thirty pounds.

The numeral thirty is the predicate of the nominal the boys in (6a), whereas it may not be a predicate of the nominal the gold in (6b). Comparing (6b) and (6c), we see that the numeral needs the support of the measure word pounds to function as the predicate of the string the gold. Following the assumption that the copula in a nominal predicate construction is a tense-bearer and the surface order of the subject-copula string is derived by the raising of the subject from its base-position, I assume that in (6a), the subject nominal is merged with the numeral predicate directly in its base-position, whereas this is impossible in (6b). The contrast is related to the nominal type of boy and that of gold.

I use the feature ‘numerability’ to represent the contrast between nominals that may combine with a numeral directly and nominals that may not do so. Accordingly, [+ numerable] means allowing a numeral, and [− numerable] means disallowing a numeral. Therefore, the nominals in (4) and (6a) are [+ numerable] and those in (5) and (6b) are [− numerable].

The numerals in the nominals in (4), which are all [+ numerable], are different. In this analysis, numerability cares about the ability to occur with a numeral only,
and no special status is given to the contrast between singularity and plurality, or among integer, zero, and other numerals.

The feature of numerability is attested in the fact that certain elements intrinsically bring about a relevant effect. In Dutch, the presence of a collective affix such as *werk makes the noun unable to occur with any numeral (de Belder 2011a: 218) and thus the affix is a marker of [−numerable] in my analysis. In (7a), the nominal suiker ‘sugar’ has [+numerable], since it occurs with the numeral drie ‘three’. In both (7b) and (7c), *werk occurs. In the presence of the numeral drie ‘three’, (7b) is not acceptable. The acceptability contrast indicates that it is the suffix that brings about the feature [−numerable] to the nominal (col = collective).

(7) a. drie suiker-en b. *drie suiker-werk-en
   three sugar-pl. three sugar-col-pl.
   ‘three sugars’
   c. suiker-werk
      sugar-col
      ‘confectionery’

In addition to numerability, we also identify the feature ‘dimensionality’. Some words may be modified by a size-denoting expression (e.g. big, small) or shape-denoting expression (e.g. long, round, square, thin) and some may not. Following Dixon (1982), I call the two types of expressions dimension expressions. In (8a), (8b), and (8c), the dimension adjectives big, large, and square modify the nouns apple, furniture, and watermelon, respectively. In (8d), (8e), (8f), and (8g), however, large, big, and square may not modify oil, music, noun, and wine (see Jespersen 1924: 198, Quine 1960: 104, McCawley 1979 [1975]: 170, Bunt 1985a: 199).

(8) a. a big apple b. large furniture c. square watermelon
d. *large oil e. *large music f. *big noun
g. *square wine

The contrast is also found in predication, as seen in (9) (from Chierchia 2010: 110):

(9) a. The violets are small.
    b. The furniture is small.
    c. *The snow is small.

In (9a), the dimension adjective small is the predicate of the violets. Similarly, in (9b), the adjective is the predicate of the furniture. In (9c), however, the adjective may not be the predicate of the snow.

I use the feature ‘dimensionality’ to represent the contrast between nominals that may be modified by a dimension modifier and nominals that may not be. Thus, [+dimension] means allowing a dimension modifier, and [−dimension] means
disallowing a dimension modifier. The nominals in (8a), (8b), (8c), (9a), and (9b) are [+ dimension] and other nominals in (8) and (9) are [− dimension].

When a nominal has [+ dimension], we know that its denotation must have ‘a certain shape or precise limits’ (Jespersen 1924: 198). The shape or limits are definable or measurable in certain dimensions (e.g. length, size, volume and shape), and therefore, atomicity is exhibited. In contrast, a nominal with [− dimension] denotes either material, which in itself is independent of shape or size, such as *silver, water, butter, gas, air*, or immaterial notions, such as leisure, music, traffic, success, tact, *common sense* (Jespersen 1924: 198). In my understanding, the former group of nouns can occur with a standard or container measure, as seen in (10a) and (11a), whereas the latter group cannot, as seen in the rest of the examples in (10) and (11).

(10)  a. a kilo of butter   b. *a kilo of leisure   c. *a kilo of nouns

(11)  a. a bowl of butter  b. *a bowl of leisure  c. *a bowl of nouns

Similar to numerability, dimensionality is also attested in the fact that certain elements intrinsically bring about a relevant effect. For instance, shui ‘water’ alone may not be modified by xiao ‘small’, as seen in (12a); but if it is followed by a CL such as di, the whole compound shui-di can be modified by xiao, as seen in (12a’). Similarly, ni ‘mud’ alone may not be modified by xiao, as seen in (12b); but if it is followed by a CL such as kuai, the whole compound ni-kuai can be modified by xiao, as seen in (12b’) (this issue is further discussed in Zhang 2011b).

(12)  a. *xiao shui  a’. xiao shui-di
      small water         small water-CL
      ‘small drop(s) of water’

      b. *xiao ni   b’. xiao ni-kuai
      small mud          small mud-CL
      ‘small chunk(s) of mud’

      c. *da yun  c’. da yun-duo
      big cloud         big cloud-CL
      ‘big piece(s) of cloud’

It is important to point out that words such as big and small and their Chinese counterparts have an intensifying usage. As stated in Morzycki (2009: 176), ‘an adjective that normally expresses size characterizes the degree to which the gradable predicate holds’, as shown in the examples in (13).

(13)  a. big idiot           b. big smoker
      c. da hao xingshi  d. xiao xian shenshou
      big good situation small show skill
      ‘very good situation’    ‘show the skill a little bit’
The intensifying readings are not size readings, and thus the adjectives in such a use are not dimension adjectives.

12.2.2 Defining count and mass by the two features

Traditionally, the notion of count is in direct contrast to the notion of mass. Different from this binary analysis, I use the two values of the two features, numerability and dimensionality, to describe the count-mass contrast.

The feature numerability alone may distinguish a count noun from a non-count noun. If a nominal may combine with a numeral directly in the context, it has [+numerable] and thus is a count nominal in that context. Otherwise, it is a non-count one. According to Chierchia (1998a: 353, 2010: 104), being able to combine with a numeral is the signature property of a count nominal.

But numerability alone is not enough to identify whether a noun is a mass noun. A non-count noun is not necessarily a mass noun. On the one hand, well-recognized mass nouns, such as the word oil, may be neither combined with a numeral directly (see (5a)), nor modified by a dimension adjective (see (8d)). On the other hand, words such as furniture may be modified by a dimension modifier (see (8b)), although they cannot combine with a numeral directly (see (5b)). Such nouns are non-count and non-mass. I claim that although the feature [+numerable] alone is enough to define the count status of a nominal, it is the combination of both [+numerable] and [−dimension] that defines the mass status of a nominal.

The four possible combinations of the two values of the features are summarized in Table 12.1:

<table>
<thead>
<tr>
<th>Numerable</th>
<th>Dimension</th>
<th>Example</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>+</td>
<td>+</td>
<td>apple in (4a), (8a)</td>
<td>count</td>
</tr>
<tr>
<td>+</td>
<td>−</td>
<td>noun in (4f), (8f)</td>
<td>count</td>
</tr>
<tr>
<td>−</td>
<td>+</td>
<td>furniture in (5b), (8b)</td>
<td>non-count, non-mass</td>
</tr>
<tr>
<td>−</td>
<td>−</td>
<td>oil in (5a), (8d)</td>
<td>mass</td>
</tr>
</tbody>
</table>

Among the four possibilities in Table 12.1, (a) and (b) are both count, (d) is mass, and (c) is non-count and non-mass. Although count is not mass and mass is not count, what is new in this analysis is the independent status of (c). The existence of this group of nouns indicates that non-count nominals do not have to be mass nouns. Also, from a different perspective, having the feature [−dimension] means that the noun is not a mass noun, but it does not mean that the noun must be a count noun (contra Wiltschko 2005a, among others). *Duckling* and the German word *Eichhörnchen* 'squirrel' can occur as non-count nouns, in addition to count
nouns (see de Belder 2011b, fn. 13), although they can be modified by dimension modifiers (e.g. small duckling). Moreover, the independent status of (b) indicates that not all count nouns denote entities that have physical dimensions, since not all count nouns may be modified by a dimension modifier.

I claim that the two features, numerability and dimensionality, are universal in defining count and mass nouns, assuming numerals and dimension modifiers are available in all languages. Also, they are the only criteria to be considered in analyzing the count-mass contrast.

12.2.3 Attesting the two features in co-occurrence restrictions

The linguistic reality of numerability and dimensionality is independently attested in co-occurrence restrictions of articles, quantifiers, and classifiers (CLs).

It is well-known that indefinite articles and some quantifiers occur with count nouns in English. For instance, every and many occur with nouns that have [+numerable], and much occurs with nouns that have [−numerable].

(14)  
  a. {every} apple   
  b. {many/*much} apples 

(15)  
  a. *{every} oil         
  b. {*many/much} oil     
  c. {*many/much} furniture

In Chinese, some CLs are sensitive to the dimension feature of the noun. Some CLs take nouns with [−dimension] only. For instance, no liquid-denoting noun may be modified by a dimension adjective, as seen in (16a). Such a noun is [−dimension]. It can occur with the CL di, as seen in (16b).

(16)  
  a. *chang {you/shui/xue/niao/yanlei}   
       long oil/water/blood/urine/tear  
  b. san di {you/shui/xue/niao/yanlei/*putao}   
       three oil/water/blood/urine/tear/grape  
       ‘three drops of {oil/water/blood/urine/tear/*grape}’

In contrast, putao ‘grape’ can be modified by a dimension adjective, as seen in (17a) below. Such a noun is [+dimension]. It may not occur with di, as seen in (16b) above. Other CLs that reject nominals with [−dimension] include ji (for liquid medicine), pao (for urine), and tan (for any liquid). I call such CLs (part of Chao’s 1968 partitive CLs) individuating CLs, which select [−dimension].

(17)  
  a. da putao  
       big grape  
       ‘big grape’   
  b. san ke {putao/*you/*xue/*yanlei}  
       three grape/oil//blood/tear
Words like *putao* can be selected by another kind of CL, called ‘individual CLs’ in Chao (1968). The CL *ke* in (17b) is such a CL. It selects nouns with [+dimension]. Other CLs such as *ben*, *tou*, and *zhi* are also individual CLs. Moreover, collective CLs, such as *zu* ‘group’, *qun* ‘crowd’, *da* ‘dozen’, *shuang* ‘pair’, and *dui* ‘pair’, and partitive CLs, such as *ye* ‘page’, *duan* ‘paragraph’, and *zhang* ‘chapter’ (they are also Chao’s 1968 partitive CLs), also occur with nouns with [+dimension].

12.3 The two features in nouns

12.3.1 Numerability of nouns

Occasionally, we see people claim that numeral CL languages do have count nouns, or people feel reluctant to admit that there is no count noun in such languages. However, we still need to consider ‘[i]f we assume that classifier languages have count nouns (similar to English *silverware*, cf. constructions like *three pieces of silverware*), then it is unclear what necessitates the use of classifiers’ (Krifka 2008: 5).

If we put unit words such as *dui* ‘pile’ aside, it is undeniable that no noun in Chinese is able to combine with a numeral directly, as shown in (18). Therefore, all nouns in the language have the feature [−numerable]. This means that no noun in the language is a count noun.

(18) a. *san* xianglian  b. *san* you
three  necklace  three  oil

The occurrence of a unit word such as a CL is obligatory between a numeral and a noun in Chinese. In contrast to Chinese, in some languages such as Hopi (Whorf 1956 [1941]: 141, Greenberg 1990 [1972]: 176), Halkomelem Salish (Wilhelm 2008: 64), and Yudja (an indigenous language spoken in Brazil; see Lima 2010b), all nouns can combine with a numeral unconditionally. (19) is an example (Lima 2010b: 7).

(19) txabïa apeta  [Yudja]
three  blood
‘three units of blood’
(the unit is identified in the context: drops, puddles, or containers)

In Yudja, there are neither CLs nor plural markers. When a numeral and a noun are combined, the exact unit of counting depends on the discourse context (Lima 2010b: 13). Lima reports the naturalness of data like (19) in the absence of either Universal Sorter or Universal Packager effects. We can see that all nouns in such languages may have the feature [+numerable].

Between the above two patterns, in languages such as English, some nouns may combine with numerals directly (e.g. *apple*) and some may not (e.g. *oil*), as seen before.
Dimensionality of nouns

Although all nouns in Chinese are non-count nouns, they are not all the same with respect to dimensionality. In 12.2.3, we see that nouns with [+dimension] are selected by individual CLs, and nouns with [−dimension] are selected by individuating CLs. As shown in (20), nouns such as he ‘river’ can be modified by a dimension adjective such as changchang ‘long’. In contrast, material nouns such as you ‘oil’ in (21a) and immaterial nouns such as minzhu ‘democracy’ in (22a), reject such adjectives (putting aside the intensifying reading of such adjectives; see the discussion of (13)).

(20) a. changchang de he  b. da qi-qiu  c. fang xigua
     long  de  river  big  air-ball  square  watermelon
     ‘long river’  ‘big balloon’  ‘square watermelon’

(21) a. *changchang (de) you  b. *da (de) zheng-qi
     long  de  oil  big  de  steam-ai
     c. *fang  de  mianfen
     square  de  flour

(22) a. *changchang (de) minzhu  b. *bo (de) zibenzhuyi
     long  de  democracy  thin  de  capitalism

The constraint is shown not only in modification, but also in predication. The string hen chang ‘very long’ may not be the predicate of the mass noun you ‘oil’ in (23a), but it can be the predicate of the non-mass noun he ‘river’ in (24a).

The string hen da ‘very big’ may not be the predicate of the mass noun zheng-qi ‘steam-air’ in (23b), but it can be the predicate of the non-mass noun qi-qiu ‘balloon’ in (24b).

     oil  very  long  steam-air  very  big

(24) a. He  hen  chang.  b. Qi-qiou  hen  da.
     river  very  long  air-ball  very  big
     ‘The river is very long.’  ‘The balloon is very big.’

This contrast shows that the feature dimensionality can divide Chinese non-count nouns into the mass-type, which has [−dimension], and the non-mass-type, which has [+dimension].

Greenberg (1972: 26) claims that nouns in CL languages have the characteristics of a mass noun. The idea is also seen in Hansen (1972), Krifka (1995), Doetjes (1996), Chierchia (1998a), among many others. According to our new analysis of the count-mass contrast, however, not all nouns in Chinese are mass nouns.
12.4 The two features in unit words

12.4.1 Classification of unit words

All unit words tell us what counts as one in counting. Unit words include CLs and measure words. The latter group is composed of standard measures such as *kilo* and container measures such as *cup* in *three cups of tea*.

Among various types of CLs, kind CLs have no occurrence restrictions. They occur with all types of nouns. Standard and container measures occur with either [+dimension] nouns or material type of [−dimension] nouns, but reject nouns denoting immaterial notions (see (10) and (11)). So these three types of unit words are not sensitive to the contrast between [+dimension] and [−dimension]. In the following data, the nouns in the a-examples are [+dimension] and those in the b-examples are [−dimension].

(25) a. shi zhong luobo b. shi zhong mianfen [kind cl]
ten cl. carrot ten cl. flour
‘ten types of carrot’ ‘ten types of flour’

(26) a. shi gongjin luobo b. shi gongjin mianfen [standard measure]
ten kilo carrot ten kilo flour
‘ten kilos of carrots’ ‘ten kilos of flour’

(27) a. shi xiang luobo b. shi xiang mianfen [container measure]
ten box carrot ten box flour
‘ten boxes of carrots’ ‘ten boxes of flour’

When these three types of unit words occur with nouns of [+dimension], they do not represent the natural units of the elements encoded by the nouns.

Unit words that select [−dimension] nouns are individuating CLs (section 12.2.3), as shown in (28). Such CLs occur with mass nouns (e.g. Croft 1994: 162). Semantically, individuating CLs are associated with the idea that the noun refers to some kind of mass and the CL gives a unit of this mass (Aikhenvald 2003: 318).

(28) a. shi dui tu b. wu gu zheng-qi
ten cl. earth five cl. steam-air
‘ten piles of earth’ ‘five puffs of steam’

c. wu zhang zhi d. wu di you
five cl. paper five cl. oil
‘five pieces of paper’ ‘five drops of oil’

e. wu tan you f. wu pao niao
five cl. oil five cl. urine
‘five puddles of oil’ ‘five units of urine’


Unit words that occur with [+ dimension] nouns are divided into three types: (1) What counts as one is bigger than the natural unit of the element denoted by the non-mass noun. In this case, a collective CL is used, as in (29a). Collective CLs (called group measures in Chao 1968: 595) include the so-called arrangement CLs, such as pai ‘row’ and luo ‘stack’, and number set CLs, such as shuang ‘pair’, dui ‘pair’, and da ‘dozen’. (2) What counts as one is smaller than the natural unit. In this case, a partitive CL is used, as in (29b). (3) What counts as one matches the natural unit. In this case, an individual CL is used, as in (29c).

(29) a. shi dui luobo
   ten cl carrot
   ‘ten piles of carrots’

b. shi pian luobo
   ten cl carrot
   ‘ten slices of carrot’

c. shi gen luobo
   ten cl carrot
   ‘ten carrots’

Generally speaking, the same form of a unit word can belong to different types, depending on the type of the associated noun, and the semantic function of the unit. In (28a), the CL dui occurs with the mass noun tu ‘earth’, and it is thus an individuating CL. However, in (29a), dui occurs with the non-mass noun luobo ‘carrot’, and it is thus a collective CL.

12.4.2 Unit words as the unique numerability bearers in Chinese

All unit words may combine with a numeral directly, although under certain conditions, the numeral yi ‘one’ can be implicit.

The conditions of silent yi are studied by Yang (1996). Silent yi can follow mei ‘each’ and the demonstrative zhe ‘this’ or na ‘that’. When yi is covert, its occurrence in syntax can be attested by the singular reading of the whole nominal.

(30) Shufen xiang mai zhe (yi) ben shu.
   Shufen want buy this one cl book
   ‘Shufen wants to buy this book.’

Therefore all unit words have the feature [+ numerable].
Since no noun in Chinese has the feature \([ + \text{numerable}]\) and all unit words have the feature, the latter are the unique numerability bearer in the language. We can see that numerability does not have to be anchored to lexical or root elements (for a non-lexical approach to the count-mass contrast, see Allan 1980, Borer 2005, Pelletier 2009; for a lexical approach to the contrast, see Doetjes, to appear).

The analytical realization of numerability is parallel to the situation that tense information can be realized by either verbs or auxiliaries in English. Formal features in general can be distributed in various types of elements.

### 12.5 Comparing with the dichotomous-contrast analysis

The study of the contrast between count and mass nouns dates back to Aristotle. Developing the insights of many previous studies, I have made the following two main claims with respect to the contrast.

First, a count noun is defined exclusively by \([ + \text{numerable}]\), i.e. the possibility to combine with a numeral directly. It has been generally recognized that such a combination possibility is the ‘signature’ grammatical property of count nouns (e.g. Chierchia 2010: 104). I have now further argued that this is the only defining grammatical property of a count noun, cross-linguistically. This syntagmatic definition means that the count/non-count distinction is clearly linguistic, rather than extra-linguistic. It is thus not surprising that countability is expressed in various ways, cross-linguistically and within the same language. In Chinese, generally speaking, no noun may combine with a numeral directly, and therefore, no noun is a count noun. Numerability is instead represented exclusively by unit words, including CLs and measure words. In languages such as Yudja (Lima 2010b) and Halkomelem Salish (Wilhelm 2008: 64), no CL exists, and every noun can combine with a numeral directly. Thus every noun can be a count noun. Between these two patterns, in languages such as English and Dëne (Wilhelm 2008), in an unmarked situation (i.e. without a shift), some words are \([ + \text{numerable}]\), and others are \([-\text{numerable}]\).

Second, the notion of mass is not the direct negation of count. Instead, it is the combination of the two syntagmatic properties: \([-\text{numerable}]\) and \([-\text{dimension}]\). Words such as *oil* in English and their counterparts in Chinese are mass nouns. This refined analysis makes it possible to precisely identify elements that may not combine with a numeral directly but may allow a dimension adjective, e.g. *furniture* in English and *pingguo* ‘apple’ in Chinese. Such words do not denote massive objects. As Chierchia (2010: 144) put it, ‘we know right off the bat that *furniture* cannot be treated on a par with *water.’ Such words have been identified as ‘count mass nouns’ (Doetjes 1996: 44, 2010: 44), ‘object mass nouns’ (Barner & Snedeker 2005), and ‘fake mass nouns’ (Chierchia 2010: 110). The similarity between such
words and the Chinese counterparts of English count nouns has been mentioned in Doetjes (1996: 35), Krifka (2008: section 6.5), Cowper and Hall (2009a: 1), and Chierchia (2010: 111, fn. 12), among others. In Doetjes (1996: 34), ‘count mass nouns’ are the words which are semantically count but behave like a mass noun syntactically. If one assumes that there is a binary contrast and then gives a name to the elements that do not fit in the contrast, I do not think the analysis is complete.

Although it has been widely believed that all nouns are mass nouns in Chinese, the difference between English typical mass nouns and Chinese non-mass nouns, with respect to dimensionality, has been noted in Gil (2008: 8). He finds that unlike the former, the latter can be modified by ‘size and shape adjectives’. In my analysis, both furniture and pingguo are [−numerable] and [+dimension]. Therefore, da pingguo ‘big apple’ is as natural as big furniture.

The close interaction between the notions numerability and dimensionality has long been realized in the literature, but the nature of the relation between them has not been clarified (see Jespersen 1924: 198). Quine (1960: 104) notes that unacceptability of *spherical water and *spherical wine. On the other hand, it is obvious that count nouns such as suggestion also reject spherical. Bunt (1985a: 199) also points out that mass nouns such as water may not be modified by adjectives such as large. However, McCawley (1975: 170) finds that furniture and footwear, which have also been treated as ‘mass’ nouns, admit size modification ‘much more readily than hard-core mass nouns such as rice’.

The most recent and thorough discussion of the relation between the count-mass contrast and size adjectives is de Belder (2011a,b). Her discussion does not cover other dimension modifiers such as thick, thin, round, though. Crucially, she claims that ‘if something acquires the [size] feature, it automatically becomes countable.’ (2011b: 183) So for her analysis, size features entail the count status. This is different from my analysis, which gives an equal status to numerability and dimensionality: neither entails the other, and thus there are four possibilities. One empirical consequence of her analysis is that she fails to capture the fact that non-count nouns such as furniture may have the size feature. This kind of noun is predicted to be ‘illicit’ in her theory (de Belder 2011a: 83 (34), 2011b: 180), contrary to the fact.

The idea that count and mass is not a dichotomous contrast and thus we need more features to represent them is also seen in Muromatsu (2003) and Acquaviva (2010). However, in the absence of syntactic criteria to analyze the empirical issues, the idea is somewhat undeveloped. But it does pave the way for the research in this chapter.

The proposed two features, numerability and dimensionality, are different, but are both related to the countability of nominals. This is parallel to our understanding that tense and aspect features are different, but are both related to the temporal properties of linguistic elements.
12.6 Reflections on theories of the relation between CLs and countability

12.6.1 The syntactic foundations of the presence of CLs

The novel analysis of the count-mass contrast proposed in this chapter opens a new window to see the syntactic foundations of individual CLs in CL languages.

Counting is possible in the presence of a unit. The unit tells us what counts as one in the context. The general function of a unit word is to specify the unit for counting. Such a word is [+ numerable]. In Chinese, when a noun occurs with a CL, it is the CL rather than the noun that is the bearer of numerability.

Individual CLs are syntactically different from nouns. Thus the fact that numerability is realized on CLs rather than nouns is a syntactic issue. The syntactic nature of the existence of individual CLs can be seen in another fact: the occurrence of such CLs is sensitive to syntactic categories in English. Counting in verbal phrases in English requires the occurrence of CLs (Krifka 2007: 39), but not in nominals, as seen in (31). There is no CL in the nominal counting construction three trips to Paris in (31a), but the CL times is obligatory in the verbal counting construction in (31b). Like nominals in Chinese, verbal phrases in English are not numerability bearers, and thus need CLs in counting. If we consider the representation of numerability in verbal phrases, English should be treated as a CL language.

(31) a. Bill made three trips to Paris.
   b. Bill traveled to Paris three *(times).

This numerability-bearer analysis of CLs calls for a review of our current understanding of CLs in CL languages.

12.6.2 How special are the CLs of CL languages?

All seven types of unit words listed in 12.4.1 ‘are closely related in grammar and function’ (Croft 1994: 152). Like measure words, CLs are also counting units or ‘unit counters’ (Allan 1977: 293).

It has been widely believed that all nouns in CL languages are mass nouns, and therefore, the basic function of CLs is to divide mass into units (e.g. Quine 1969, Greenberg 1972: 26, Link 1991, Borer 2005: 101, Krifka 2008: section 6.3). Accordingly, the syntactic projection headed by a CL has been called DivP (Borer 2005). In the previous discussion, I have shown that CL languages can distinguish mass from non-mass nouns. As a consequence, the general function of CLs is not dividing or individuating.

Let us examine how the ‘dividing’ assumption misrepresents the basic function of CLs. We have introduced five types of CLs in 12.4.1:
(32) a. san zhong yang
   three cl sheep
   ‘three kinds of sheep’

b. san di shui
   three cl water
   ‘three drops of water’

c. san qun yang
   three cl sheep
   ‘three groups of sheep’

d. san pian xigua
   three cl watermelon
   ‘three slices of watermelon’
e. san ben shu
   three cl book
   ‘three books’

From the translations of (32a), (32b), (32c) and (32d), we can see that English also has kind CLs such as kind, individuating CLs such as di ‘drop’, collective CLs such as group, and partitive CLs such as slice.

Among the five types of CLs, the individuating CL in (32b) indeed divides a massive object into units (also see the examples in (28)). CLs like this are called ‘partitive measures’ in Chao (1968), and ‘classifiers for massive objects’ in Gerner and Bisang (2010: 606). Such CLs are also found in non-CL languages such as English, as seen in the word drop in the translation of (32b). Obviously, individuating CLs cannot distinguish CL languages from other languages.

What English does not have is individual CLs. There is no English counterpart for ben in (32e). It is this type of CL that distinguishes CL languages from non-CL languages such as English. In non-CL languages, individual CLs are not overtly represented by linguistic expressions. In such languages, it has been assumed (Quine 1969: 36) that the semantics of an individual CL is integrated either in the numeral (see Wilhelm 2008: 55) or the noun (see Chierchia 1998a).

Crucially, individual CLs do not divide or individuate anything. They do not occur with mass nouns. As pointed out by Bale and Barner (2009b: 7), ‘default classifiers [such as the individual CL ge in Mandarin Chinese] often combine with nouns that already are interpreted as containing individuals’. Such CLs neither individuate anything nor create new units for the individuals any more, unlike collective or partitive CLs. Therefore, the popular belief that it is the individuating (discrete set-creating) function of CLs that is special in CL languages needs reconsideration. A more precise generalization is that in addition to the various ways of specifying a unit for counting, CLs in CL languages may also represent the natural unit of entities that show atomicity, whereas the CLs of other languages do not have
this semantic function. In other words, the CLs in CL languages are special in their ability to represent the natural units of the entities denoted by non-mass nouns.

If the general function of CLs is not dividing, we need to reconsider Borer’s (2005) syntactic analysis of the count-mass contrast. In her analysis, the absence of ‘dividing structure’ (DivP) derives ‘mass’ readings, and the presence of dividing structure derives ‘count’ readings. The two features proposed in this chapter call for a richer structure to represent the count-mass contrast.

12.6.3 The unreliability of the de and pre-CL adjective arguments

In this section, I argue against the assumed correlation between the count-mass contrast and two phenomena in Mandarin Chinese: the occurrence of an adjective to the left of a unit word and the occurrence of the functional word de to the right of a unit word (Cheng and Sybesma 1998, 1999).

I have argued that in Chinese, neither nouns nor CLs make a distinction between count and non-count themselves, since all nouns are non-count elements (12.3.1) and all CLs are count elements (12.4.2). But the selection of CLs may distinguish mass nouns from non-mass nouns. Individual, collective, and partitive CLs occur with non-mass nouns and individuating CLs occur with mass nouns (other unit words, i.e. kind CLs, standard and container measures, occur with both mass and non-mass nouns).

Cheng and Sybesma (1998, 1999) try to make a distinction between count CLs and mass CLs (called massifiers). The names are used to show that in Chinese, the contrast between count and mass nouns can be distinguished at the level of CLs, if not at the level of nouns.

In Cheng (2009b: 3), it seems that count CLs are equivalent to individual CLs and all other kinds of unit words are mass CLs. Cheng and Sybesma (1988, 1999) formalize the following two criteria:

Criterion A: A pre-CL adjective may occur with a mass noun, as seen in (33a), but not with a ‘count’ noun, as seen in (33b) (Cheng & Sybesma 1998: 390, 1999: 516). The term ‘count noun’ in their analysis is called non-mass and non-count noun in this chapter.

(33) a. yi da zhang zhi one big cl paper
   b. *yi da wei laoshi one big cl teacher
   c. yi da tiao hao-han good-guy three big cl tiger
   d. san da zhi laohu three big cl necklace
   e. san chang tiao xianglian necklace

   ‘one big piece of paper’
   ‘one big good guy’
   ‘three big tigers’
   ‘three long necklaces’
It is true that (33b) is not acceptable. But isolated cases like this do not affect the observation that other examples of the same type are acceptable, as shown in (33c,d,e) (see Cheng and Sybesma 1998: 390 fn. 4 for their acknowledgement of counter-examples). Tang (2005), Hsieh (2008), and Li (2011: 34), among others, all present many counter-examples to this claim about pre-CL adjectives. More examples can be found in Zhu (1982: 52), Lu (1987), and Luo (1988). Therefore, the adjective criterion is empirically problematic.

**Criterion B:** De may occur between a measure word and a mass noun, but not between a CL and a ‘count’ noun (Chao 1968: 555, 588; Zhu 1982: 51; Cheng and Sybesma 1998: 388, 1999: 515). A typical pair of examples is (34):

(34) a. san wan de tang b. *san ge de laoshi

three bowl de soup three CL de teacher

‘three bowls of soup’ ‘three teachers’

Again, the unacceptability of (34b) is one of a few isolated cases. In fact, all types of CLs can be followed by de in an appropriate context. The choice of the context has nothing to do with the count-mass contrast. Instead, it has to do with the syntactic position of de. In Zhang (2011a), I show that there are two sources of de: one introduces a constituent directly and the other surfaces in a comparative ellipsis construction. Constructions of individual, individuating, and kind CL host the latter de only, whereas those of the other types of unit words (partitive and collective CLs, standard and container measures) host de of either source. Note that the division here does not match with Cheng and Sybesma’s distinction between count and mass CLs. If one just considers the phonological form of de without considering its structural position, then, de may occur with all types of CLs or unit words, as shown in (35).

(35) Shufen chi-le yi-bai {ge/gongjin/bao/pian/dui/zhong} de pingguo.

Shufen eat-perf one-hundred CL/kilo/bag/slice/pile/kind de apple

‘Shufen ate 100 apples or 100 {kilos/bags/slices/piles/kinds} of apples.’

Therefore, Cheng and Sybesma’s (1998, 1999) claim that one type of CL (the ‘count’ type) may not be modified by an adjective, and may not be followed by de, whereas the other type (the ‘mass’ type) can, is descriptively inadequate.

As mentioned above, several works, including Tang (2005: 432, 440–446), Hsieh (2008: 34), Li (2011), etc. have already presented many counter-examples to falsify the alleged distinction. Wu and Bodomo (2009: 489) point out that the two alleged types of CL can occur with the same NP (See also Borer 2005: 98), as shown in (36). *Ben* in (36a) and *li* in (36b) are count CLs, and *xiang* in (36a’) and *wan* in (36b’) are mass CLs, in Cheng and Sybesma’s system.
Her and Hsieh’s (2010: 541) following examples show that the two constraints on the so-called count CLs (i.e. individual CLs) can even be violated at the same time. The CLs ke in (37a) and tiao in (37b) are typical individual CLs, but they are both preceded by a modifier and followed by de.

In conclusion, the alleged two criteria cannot make any distinction in CLs in Chinese, regardless of whether the assumed distinction correlates with the count-mass contrast.

12.7 Summary

Based on Mandarin Chinese, in this chapter I have argued for a more refined syntactic analysis of the count-mass contrast. I list my main conclusions as follows.

The count-mass contrast of linguistic elements is decomposed into two features: [numerable] and [dimension]. [+numerable] means a noun can combine with a numeral directly, and thus it is a count noun. [−numerable] nouns are non-count nouns. Nouns in Chinese are non-count nouns in general. [+dimension] means a noun can combine with a shape or size modifier, and thus it is a non-mass noun. A mass noun is defined by both [−numerable] and [−dimension]. Not all nouns in Chinese are mass nouns.

Some CLs select mass nouns and some select non-mass nouns. The latter type of CL, i.e. individual CLs, distinguishes CL languages from other languages. Like other types of unit words, all CLs specify units and tell us what counts as one in counting. The semantic function of individual CLs is to represent the natural units of the elements denoted by non-mass nouns. Such a CL has no dividing function at all.

The feature [+numerable] may be distributed in different types of elements cross-linguistically and within the same language. In Mandarin Chinese, nouns do not have this feature, whereas unit words do. In languages such as English, the feature is found in both count nouns and unit words such as measure words.
Semantic triggers, linguistic variation and the mass-count distinction*

ALAN C. BALE AND DAVID BARNER

13.1 Introduction

Although most languages allow nouns to be used with numerals to express cardinality, they differ significantly in how they grammatically encode such expressions. Some languages, like English, require count syntax whereas others, like Mandarin, lack count syntax and typically use classifiers. Here, we ask what appears to be a simple question: how do children determine whether their language makes a distinction between mass and count syntax? This question reveals itself to be subtle and difficult when languages beyond English and Mandarin are considered. We argue that prototypical syntactic and morphological differences between mass-count and classifier languages are not constitutive of this typological difference. The use of classifiers, the combination of numerals with bare nouns, and even plural morphology can occur in both mass-count and classifier languages. As a result, such features cannot be sufficient for determining whether or not a language has count syntax. Instead, we argue that it is the relation of these syntactic structures to their semantic interpretations that differentiates languages and guides acquisition. Only mass-count languages can specify exclusive reference to singularities in the absence of classifiers or measure words.

This proposal is laid out in three sections. Section 13.2 reviews the syntactic and semantic differences between mass-count languages and classifier languages, using English and Mandarin as a starting point. Our main conclusion, following many others, is that English makes a mass-count distinction that is not paralleled in

* We would like to thank Peggy Li, Yaling Hsiao and Hrayr Khanjian for thoughtful and valuable comments on this paper and, in particular, its linguistic examples. All errors are due to the authors. We would also like to thank Pierina Cheung for her contributions to the data analysis as well as two anonymous reviewers for their comments and suggestions on previous versions of this paper. This work would not have been possible without a standard research grant to Alan Bale from Social Sciences and Humanities Research Council of Canada (SSHRC) (“Language, Number and Mathematics”) and a Scholar Award to David Barner from the James S. McDonnell Foundation.
classifier languages like Mandarin. Unlike classifiers, the presence or absence of count syntax in English fundamentally shifts the interpretation of individual nouns, whether known or novel. However, despite this difference, Mandarin and English both allow non-count nouns to denote sets of individuals. In Mandarin, any noun that has a count-noun counterpart in English (e.g. cup) can be used with or without classifiers to refer to sets of individuals. In English, many mass nouns (like furniture and equipment) denote individuals. Critically, however, both types of nouns contrast with singular count nouns, which are exclusive to English, and are typically used to talk about single individuals.

Having outlined the cross-linguistic differences between mass-count and classifier languages, section 13.3 reviews the evidence of when children acquire count syntax, plural morphology, number words, and classifiers cross-linguistically.

Finally, section 13.4 explores the evidence that children might potentially use to determine whether their language makes a mass-count distinction. First, we consider two syntactic cues (pluralization and classifier usage), and argue that each fails to explain how children differentiate mass-count languages and classifier languages early in acquisition. Second, we argue that children are sensitive to semantic differences between nouns, and in particular can use singular reference to infer the presence of count nouns. Our hypothesis is that consistent reference to singularities supports an inductive inference in English that some noun phrases with singular agreement (e.g. those that will eventually be acquired as singular count nouns) can only denote singleton sets. In other languages, no such consistent evidence is provided for non-plural nouns, and thus no induction regarding singular-plural status is supported. This difference, we argue, is the basis for children’s inference that languages like English have count syntax, whereas languages like Mandarin, Armenian, and others do not.

13.2 English, Mandarin, and the mass-count distinction

Mass-count languages like English are typically distinguished from classifier languages according to a cluster of morphological and syntactic differences. In English, count nouns like chair, in contrast to mass nouns like furniture, demonstrate a singular vs. plural contrast, can occur directly with numerals, and can be modified by quasi-cardinal and distributive quantifiers (1).

(1) a. a chair / some chairs
   b. one chair / several chairs / each chair
   c. *a furniture / *some furnitures
   d. *one furniture / *several furniture(s) / *each furniture

Furthermore, mass nouns can be used as the complement of measure words in their bare, singular form whereas count nouns require plural marking (2).
Similarly, there are certain quantifiers, such as the comparative quantifiers, that require plural marking for count nouns but not mass-nouns (3).

(3) a. more furniture / *more furnitures
b. more chairs / *more chair

As demonstrated by the contrasts between furniture and chair in (1) through (3), these distinctions in English are not a simple reflection of a nominal ontology (individuated/divided denotations versus homogeneous/substance-like denotations). Although the things being talked about and referred to by furniture and chair are relatively similar (individual items of furniture), their syntactic distribution is not.

In contrast to English, classifier languages like Mandarin lack many of these distinctions. Most nouns in Mandarin cannot co-occur directly with number words. Rather they require classifiers, words akin to English measure words like piece of and grain of, as in (4).

(4) a. liang li mi / *liang mi
two cl rice / two rice
‘two grains of rice’
b. liang ge haizi / *liang haizi
two cl children / two children
‘two children’

Unlike English, Mandarin has a rich system of such classifiers which can be divided into sortal (as in 4) versus mensural, or non-sortal, subcategories (ping ‘bottle’, bei ‘cup’, dai ‘bag’, xiang ‘box’, etc.). See Tang (2005), Cheng and Sybesma (1999), and Senft (2000) for a discussion.

Another difference between English and Mandarin concerns pluralization. Mandarin does not make an obligatory singular/plural distinction. In rare cases where plural marking occurs (or pseudo plural marking as it is sometimes called in the literature—see Iljic 1994, Li 1999, Cheng and Sybesma 1999, Tang 2004, 2005), there is no singular counterpart. Nouns that lack plural marking can also refer to plural sets. In addition, there are no quasi-cardinal quantifiers in Mandarin that apply exclusively to one class of noun. Thus, Mandarin nouns exhibit many of the hallmarks of English mass nouns.

Despite these differences, it is important to note that Mandarin does make some distributional distinctions that reflect a nominal ontology. For example, certain classifiers, like ge in (4), require that their nominal complements have individuated
denotations—in other words, that the complement nouns can be used to refer to individuals—see Cheng and Sybesma (1999). Unlike in English, this distributional difference is tied closely to the nature of the nominal denotation. There are no minimal pairs like furniture vs. chair in Mandarin: i.e. pairs of nouns that can both be used to refer to individuals but have distinct syntactic distributions. Sometimes languages that have these ontologically-based distinctions have been said to demonstrate a mass-count contrast—see Cheng and Sybesma (1999), Wilhelm (2008). We distinguish between languages like Mandarin that encode the ontological distinctions from languages like English which encode a syntactic distinction that crosses ontological boundaries. It is the latter that we are calling mass-count languages.

In the following sections we draw attention to semantic distinctions between these two syntactic systems that are relevant to acquisition. One difference is that count syntax has a unique semantic property, namely that it triggers coercions from substance-like denotations to individuated denotations. A second difference concerns bare nominals. Count syntax in English includes bare nominals with singular denotations (e.g. sets of individuals such as {a, b, c}) as well as nominals with number neutral denotations (e.g. sets of individuals and groups such as {a, b, c, ab, ac, bc, abc}). In contrast, all bare nouns in Mandarin are number neutral. After noting these differences, we highlight a key similarity between the two language types: a parallelism between English object-mass nouns like furniture and Mandarin bare nouns.

13.2.1 Count syntax and coercion

Despite the syntactic differences listed at the beginning of section 13.2, some researchers have suggested that the contrast between English and Mandarin is merely a phonological illusion, and that count syntax and classifiers are represented by the same functional head (cf. Krifka 1995, Cheng and Sybesma 1999, Borer 2005). However, there is an interesting distinction between classifier syntax and count syntax that is unexplained by this hypothesis. Specifically, in English, count syntax does not just signal individuation; it triggers it grammatically. The absence of count syntax often results in an unindividuated or substance-like interpretation. In contrast, classifier syntax (at least default/sortal classifier syntax) is not necessary for individuation in Mandarin, and its removal does not result in substance-like interpretations.

In previous work, we have argued that comparative judgements reveal predictable semantic differences between nouns used in mass and count syntax in English, a method which also reveals differences between count syntax and Mandarin classifiers (Bale and Barner 2009a, Barner and Snedeker 2005, 2006). Consider the sentences in (5–8).

(5) a. Mary has more bananas than Jane does.
   b. Mary has more banana than Jane does.
a. Mary has more water than Jane does.
b. Mary has more waters than Jane does.

a. Mary has more rocks than Jane does
b. Mary has more rock than Jane does.

a. Mary has more blickets than Jane does.
b. Mary has more blicket than Jane does.

As shown in (5), nouns that are typically used in count syntax, like banana, always denote individuals as count nouns, but never when used in mass syntax. As a result, the sentence in (5a) is evaluated on the basis of number, whereas (5b) is evaluated on the basis of mass or volume. Similarly, in (6) the noun water, which is typically used as a mass noun can only denote individuals when used in count syntax (6b). Again, the decision of ‘who has more’ is based on number for the count usage (e.g. number of portions) but on volume for the mass usage. Mass-count flexible words like rock (7) also shift their interpretation as a function of mass-count syntax. Other examples include stone, paper, string, chocolate, idea, hope, thought, etc. Finally, in (8) novel words like blicket shift interpretation according to their syntactic frame (Barner and Snedeker 2006). In general, all English nouns used in count syntax denote individuals. Also, nouns that can be used in count syntax to denote individuals can never be used as mass nouns to denote countable things (for a review, see Bale and Barner 2009a; for experimental evidence in adults and children, see Barner and Snedeker 2005).

In contrast, manipulating the presence or absence of classifiers in Mandarin causes no such shift in interpretation. As noted by Krifka (1995), Cheng and Sybesma (1999), and Borer (2005), the classifiers that are closest in meaning to count syntax are the so-called sortal/default classifiers such as ge in (9).

a. liang ge haizi
two cl child
‘two children’
b. liang ge pingguo
two cl apple
‘two apples’

a. liang li mi
two cl (grain) rice
‘two grains of rice’
b. liang ba mi
two cl (handful) rice
‘two handfuls of rice’
Unlike mensural classifiers which specify how a noun’s referent should be portioned and counted (like English measure words)—see examples in (10)—default classifiers only specify that the noun’s referent can be counted, the method of individuation being determined by the content of the noun itself. However, unlike count syntax, the absence of these default classifiers does not force a noun into a substance-like interpretation. For example, the sentences in (11) are used to talk about individual children and whole apples rather than their corresponding substances.

(11) a. Haizi pao (le).
   child run (asp)
   ‘The children ran.’

b. Zhangsan de pingguo.
   Zhangsan ps apple
   ‘Zhangsan has apples.’

The lack of coercion is true for comparative constructions as well. Consider the sentences in (12).

(12) a. Zhangsan de pingguo bi Lisi de pingguo duo.
   Zhangsan ps apple than Lisi ps apple more
   ‘Zhangsan has more apples than Lisi.’

b. Zhangsan de haizi bi Lisi de haizi duo.
   Zhangsan ps child than Lisi ps child more
   ‘Zhangsan has more children than Lisi.’

Even when classifiers are not present, the nouns pingguo ‘apple’ and haizi ‘child’ denote individuals. As a result, the comparatives in (12) are evaluated by number, not mass or volume.³ Thus, Zhangsan has more pingguo than Lisi if she has a greater number of apples, even if Zhangsan’s apples are very tiny and overall weigh less (or take up less space) than Lisi’s.

Experimental evidence collected in our labs corroborates this intuition. First, when Japanese-speaking adults are asked to make comparative judgements for sentences containing translations of English count nouns (e.g. cup), their judgements do not differ from those of English speakers, and are based exclusively on number (Barner, Inagaki, and Li 2009, Inagaki and Barner, 2009). Similarly, Japanese and English speakers do not differ when interpreting nouns that denote substances (e.g. butter). Finally, when tested with translations of nouns that are mass-count

³ In most contexts, (12a) is evaluated by number. There are some contexts where an evaluation via mass is possible, such as when the apples are sliced into pieces and hence the individuals are de-emphasized (literally destroyed). In general, when the individuals are not de-emphasized, the comparative prefers an evaluation by number. In the special circumstances, it is possible that a coercion operator derives the correct interpretation of the noun.
flexible in English (e.g. *string, stone*), Japanese speakers base their judgements on number about 50 percent of the time, exactly half-way between the English results for mass and count judgements for the same types of words. In Mandarin, similar results are found (Cheung, Barner, and Li 2010, Li, Cheung, and Barner under review). Note that adding classifiers has no effect on interpretation, except for words whose counterparts are mass-count flexible in English (presumably because the classifier disambiguates between two pre-existing lexical meanings, i.e. individual vs. non-individual; see Cheung, Barner, and Li 2010, Li, Cheung, and Barner under review).

13.2.2 Count syntax and singularity

There is a contrast between two types of bare nouns in English. This contrast, which involves semantic number, does not exist in Mandarin. As discussed in Gillon (1992, 1999) and Chierchia (1998a, 2010) the denotations of bare count nouns typically contain individuals but not groups (i.e. no pluralities), whereas the denotations of mass nouns can contain both individuals and groups. For example, consider the contrast between the count noun *boy* and the mass noun *furniture* when they appear in predicate position as in (13) and (14).

(13)  a. John is a boy.
    b. Brad is a boy.
    c. #John and Brad are a boy.

(14)  a. This chair is furniture.
    b. That table is furniture.
    c. That table and this chair are furniture.

When (13a) and (13b) are true, it is extremely odd to assert (13c). In contrast, (14a) and (14b) together entail (14c). This pattern demonstrates that the denotations of singular count nouns do not contain groups whereas the denotations of mass nouns do. In other words, mass nouns are unspecified for number (see Bale, Gagnon, and Khanjian 2010, and Wilhelm 2008, for a discussion of this point; see also Corbett’s discussion of General Number in Corbett 2000).

Similar observations can be made for definite descriptions. For example, the sentence in (15a) presupposes that there is only one boy salient in the context whereas (15b) does not.

(15)  a. The boy is outside on the balcony.
    b. The furniture is outside on the balcony.

Note that, according to many speakers, the NP *the furniture* in (15b) can only be used to talk about a plurality of objects. However, this intuition does not reflect the literal meaning of *furniture*. As predicted by Gricean principles (specifically the Maxim of Quantity, Grice 1975), it would be awkward to use the superordinate word *furniture*
instead of a basic term like chair or table, especially when there is only one item being talked about. Such awkwardness can be controlled by eliminating the possible use of the basic term epistemically. For example, in a situation where one is informed that either a chair or a table is on the balcony but not both (perhaps the informant forgets which item was on the balcony), the sentence in (15b) is appropriate.

In contrast to the English examples, bare nouns in Mandarin are always number neutral (see Krifka 1995, Tang 2005, 2004; see also Kang 1994, for similar observations about Korean). Consider the sentences in (16) and (17).

(16) a. Zhangsan shi haizi.
    Zhangsan be child
    ‘Zhangsan is a child.’
  b. Lisi shi haizi.
    Lisi be child
    ‘Lisi is a child.’
  c. Zhangsan he Lisi shi haizi.
    Zhangsan and Lisi be child
    ‘Zhangsan and Lisi are children.’

(17) Haizi pao (le).
    child run (ASP)
    ‘The child/children ran.’

The noun haizi (child) is representative of all nouns in Mandarin that are used to talk about individuals. Like mass nouns in English, when the noun haizi is used in predicate position as in (16), the noun can be true of both individuals and groups. In fact, (16a) and (16b) entail (16c). Also like mass nouns in English, when the noun appears with a definite interpretation it can be used to talk about pluralities or groups. The sentence in (17) can be used to refer to a specific child or a specific group of children. There is no implication involving number. (Note that, unlike English, there are no definite articles in Mandarin. The definite interpretation occurs as the default interpretation when the noun appears in subject position.)

13.2.3 Mandarin bare nouns and English object-mass nouns

Despite the differences between English and Mandarin discussed in sections 13.2.1 and 13.2.2, there is at least one important similarity regarding the lexical representation of Mandarin nouns like pingguo (apples) and English object mass-nouns like furniture. As already shown in section 13.2.2, both types of noun are number neutral. Besides this shared property, they both also induce a comparison by number without requiring count syntax or classifier syntax. Furthermore, they both combine with semantically inert measure terms.
13.2.3.1 Comparison by number in the absence of count syntax Several previous reports have claimed that object-mass nouns like furniture and jewelry have atomic, countable individuals in their denotations and are represented like plural count nouns (Gillon 1992, 1999, Chierchia 1998a, Krifka 1995, Bale and Barner 2009a, Barner and Snedeker 2005, 2006, Barner, Wagner, and Snedeker 2008). Bale and Barner (2009a) demonstrated this using comparative constructions. As shown in section 13.2.1, English nouns that can appear in count syntax never denote individuals when used as mass nouns. In contrast, object-mass nouns like furniture denote individuals and can never appear in count syntax (in fact, in other languages, like French, their translations are often count nouns, e.g. meubles, vêtements, etc.). Consider the examples in (18) and (19).

(18)  
   a. Mary has more footwear than Jane does.
   b. Mary has more furniture than Jane does.
   c. Mary has more clothing than Jane does.

(19)  
   a. ??Mary has more footwea̧rs than Jane does.
   b. ??Mary has more furnitures than Jane does.
   c. ??Mary has more clothings than Jane does.

In each sentence in (18), the judgement is based on number (for experimental evidence, see Barner and Snedeker, 2005). Despite lacking count syntax, and despite the fact that many nouns in English require count syntax to denote individuals, these nouns denote individuals as mass nouns. Thus, these English mass nouns resemble nouns like haizi and pingguo in Mandarin, which also denote individuals without count or classifier syntax (as shown in section 13.2.1 with the examples in (11) and (12)).

13.2.3.2 Object-mass nouns and measure words Another similarity between Mandarin bare nouns and English object mass-nouns can be found in the use of measure words (see Gillon 2009). As noted in (9) in section 13.2.1, when nouns like haizi and pingguo combine with numerals they require the intervening classifier ge, which appears to contribute little to the interpretation and serves the syntactic function of licensing numeral use. For many nouns, including novel or low frequency words, this default classifier can be used instead of item-specific classifiers.

Such default classifiers are semantically similar to English measure words like item and piece, which can also be used for a variety of nouns. Critically, they can be used to combine numerals with object-mass nouns like furniture, clothing, and equipment as shown in (20).

(20)  
   a. two items/pieces of furniture
   b. two items/pieces of clothing
   c. two items/pieces of equipment
As with ge, these measure words do not specify the nature of the units to be counted. Instead, the relevant units for counting are determined by the lexical semantics of the mass noun: namely the atomic minimal parts in its denotation. As a result, comparative judgements involving these measure words generate results that are identical to those not involving measure words. (Note that a piece of furniture cannot be any piece or part—e.g., a leg of a chair—but must correspond to whole units like chairs, tables, etc.).

13.2.3.3 Summary In summary, Mandarin nouns like haizi and pingguo and English object-mass nouns: (i) denote individuals in the absence of count or classifier syntax, (ii) can be used with semantically inert classifiers or measure words, and (iii) are underspecified for number, and thus can refer to either groups or individuals. These facts suggest two conclusions about the semantics of these words. First, they all have atomic minimal parts in their denotations—e.g., individual children or pieces of furniture. The presence of atomic minimal parts provides semantic units for counting (independent of classifiers or measure words) and guides comparative

\[
\{ \text{abc, ab, ac, bc, a, b, c} \}
\]

**Figure 13.1** Representation of the denotation of haizi in a context where a, b and c are the individual children. The grey lines represent the subaggregate relation. The elements a, b, c, are the minimal parts in the denotation since there are no other members that are subaggregates of them. Furthermore they are atomic since they do not have any overlapping parts (no subaggregates that are a part of two or more minimal parts)

\[
\{ \text{a, b, c} \}
\]

**Figure 13.2** Representation of the denotation of the count NP boy without the plural morpheme in a context where a, b and c are the individual boys. Note that the denotation does not contain any groups
judgements, inducing a comparison by number thereby explaining why they can be used with semantically inert measure words or classifiers. Second, both types of nouns have denotations that not only contain atomic minimal parts, but also contain all the groups that can be formed from these minimal parts. In other words, their denotations are complete, join semilattices. This characteristic explains why such nouns can be used to talk about individuals or any group of individuals, in contrast to English bare count nouns (see Figures 13.1 and 13.2).

13.2.4 Acquiring English vs. Mandarin

The goal of this chapter is to explain how children identify and learn to interpret count syntax when it appears in their language, i.e. to explain how children discover that their language is a mass-count language, like English, rather than a language lacking such a distinction, like Mandarin. To begin, we have provided arguments to support the claim that Mandarin does not have count syntax (though it does have classifiers and measure words which often function semantically like English measure words).

The differences between these languages have important consequences for lexical representation in acquisition. If a language lacks count syntax, then the task of acquiring words that are used to refer and talk about individuals is greatly simplified. There is only one type of noun that can do this and so only one way that such a noun can be lexically represented. In contrast, if the language the child is acquiring has count syntax, then the task becomes more complex. Words that are used to refer to individuals can be represented in two ways—as a lexical item compatible with count syntax or as an object mass-noun, which can never be used in count syntax (see Figure 13.3).

In sections 13.3 and 13.4 we explore how children identify whether their language has count syntax in early acquisition, beginning with a review of when they first become sensitive to the mass-count distinction in English. We argue that, when acquiring English, children identify the basic syntax and semantics of the mass-count distinction by around 24 months of age, and that count nouns are identified as those that grammatically specify singular reference.

13.3 English and Mandarin heuristics and the age of distinction

The earliest sign that children are sensitive to the mass-count distinction in English is their use and comprehension of singular-plural morphology and agreement by around the age of 24 months (Barner et al. 2007, Brown 1973, Cazden 1968, Clark and Nikitina 2009, Fenson et al. 1994, Kouider et al. 2006, Mervis and Johnson 1991). For example, when shown two computer screens depicting sets of novel objects, 24-month-old children are more likely to look at the screen that depicts...
a plural set when told ‘Look! There are some blickets’ than when they hear ‘Look! There is a blicket!’ At this early age, children use verb agreement to guide their interpretation, as shown by the time course of their eye gaze (which shifts reliably to the plural set immediately after the verb), and as shown by their failure to differentiate singular and plural forms used without verbal agreement—e.g. ‘Look at the blicks!’

Other quantifiers begin to emerge in child speech at around the same time. Children begin using more at 19 months, some and all at 26 months, the indefinite determiner a at 27 months, and another at 29 months (Fenson et al. 1994). Children’s comprehension of these words appears to emerge at around the same time, and is correlated with their comprehension of numerals like one, two, and three (Barner, Chow, and Yang 2009). At around 2 years of age most children can recite part of their count list, and by 2 and a half have acquired the meaning of one. Between 6 and 9 months later they learn the meaning of two, then three, and then at around 3 and a half or 4 they understand that the last word used in a counting sequence labels the cardinality of the whole set (see Wynn 1990, 1992). Children’s acquisition of quantifiers and number words is also correlated in Japanese, although number word

**Figure 13.3** The different acquisitional tasks once a child knows that he/she is acquiring a language with count syntax or without

<table>
<thead>
<tr>
<th>Mandarin Acquisition Problem</th>
<th>English Acquisition Problem</th>
</tr>
</thead>
<tbody>
<tr>
<td>Word used to refer to individuals</td>
<td>Represented as bare noun with an interpretation with atomic minimal parts (unspecified for number)</td>
</tr>
<tr>
<td>Word used to refer to non-individuals</td>
<td>Represented as bare noun with an interpretation without atomic minimal parts.</td>
</tr>
<tr>
<td>Word used to refer to individuals</td>
<td>Represented as compatible with count syntax (specified for number).</td>
</tr>
<tr>
<td>Word used to refer to non-individuals</td>
<td>Represented as mass noun with an interpretation with atomic minimal parts (unspecified for number)</td>
</tr>
<tr>
<td></td>
<td>Represented as mass noun with an interpretation without atomic minimal parts.</td>
</tr>
</tbody>
</table>
acquisition begins slightly later, perhaps owing to the relative paucity of number marking in Japanese relative to English (Barner et al. 2009, Sarnecka et al. 2007).

By at least 30 months of age children are sensitive to the mass-count distinction when learning new words. When shown a novel object with a distinctive shape and material, and taught a novel label for it, children are more likely to infer that the word denotes a kind of object when it is used in count syntax than when it is used in mass syntax (Soja 1992). Specifically, when shown two new referents—one that matches the original only in shape, and the other that matches only in substance—children are more likely to point to the shape-match when tested with count syntax than when tested with mass syntax. However, it is critical to note that 30-month-olds, 3- and 4-year-olds, and even adults exhibit categorical responses only for count syntax (i.e. systematically picking the shape match when tested with solid objects). When tested with mass syntax, subjects of all ages are more equivocal, and are equally likely to extend novel mass nouns to objects as to substances, consistent with the idea that mass syntax is neutral with respect to individuation (for review, see Barner and Snedeker 2006).

By some reports, cross-linguistic differences emerge in the interpretation of novel words between 20 and 24 months. Many studies have now documented that when English-speaking children are presented with a novel word used in mass-count ambiguous syntax—e.g. Look at the bicket!—they are much more likely to extend the word by shape than by substance (see Landau, Smith, and Jones 1988). This behaviour, however, is significantly less pronounced in children and adults who speak classifier languages like Japanese and Mandarin (Imai and Gentner 1997, Li, Dunham, and Carey 2009). For example, when shown novel objects with relatively simple shapes, English-speaking 30-month-olds are more likely than same-aged Japanese children to extend a label for the object by shape. This is likely because English-speaking children must assign a syntactic representation to the ambiguous noun, and since count nouns are more frequent in children’s language input (Samuelson and Smith 1999), they are more likely to infer that the novel word is a count noun and thus that it denotes a kind of individual (Barner, Inagaki, and Li 2009, Li and Gleitman 2002, Li, Dunham, and Carey 2009). Consistent with this hypothesis, adult bilingual speakers of Mandarin and English are more likely to extend a novel ambiguous word by shape when tested in English than when tested in Mandarin, suggesting that grammatical differences between the languages drive the differences in word learning, and thus the lexical representations that are assigned to words (Barner, Inagaki, and Li 2009).

Extending these findings, Gordon (1985) and Barner and McKeown (2005) tested how children used and interpreted novel words after hearing them in syntactically neutral environments like the one above. Critically, these studies manipulated not only the shape and substance of objects, but also the number of referents. In the Barner and McKeown study, children were shown either a single novel object or a set
of ten objects of the same kind. Children then heard the object(s) labelled with mass-count neutral syntax—*Look! This is my blicket!* Next, they were shown a plural set and asked to label it. In this study and in Gordon’s, children were more likely to use a plural form of the noun if they had first heard the word used to label a singleton set than if it had labelled a plural set. In other words, children assumed that the word was a mass noun if it was first used to refer to a plural set, since it lacked plural morphology. Consistent with this, children who labelled the plural set with a plural noun were more likely to extend the word to other objects with the same shape, whereas children who omitted the plural extended the word to objects that shared the same substance. This suggests that children used numerical cues to infer the mass-count status of words, affecting both their use of the plural and whether they extended the word by shape or substance (Barner and McKeown, 2005).

Critically, Mandarin-speaking 4-year-old children show a different pattern of results. Data collected for this chapter with Pierina Cheung indicate that Mandarin 4-year-olds showed no difference in their interpretation of nouns when they were first used to label singleton or plural sets (these children were not given a singular-plural elicitation task, since plural marking is not obligatory in Mandarin). Overall, their judgements resembled those of English-speaking children who did not use the plural. This pattern of results is consistent with the idea that only English-speaking children are required to assign a syntactic mass-count representation to ambiguous novel nouns. It also supports the idea that reference to singleton sets is, by default, mapped to count syntax barring evidence to the contrary. Whereas Mandarin-speaking children encode novel nouns identically when they label sets of one object or more than one, English-speaking children draw on both syntactic and semantic assumptions when assigning novel words their lexical representations.

In summary, children make different assumptions about nouns (both novel and known) based on the type of language they have acquired. English children are sensitive to count syntax by 24 months and seem to understand the semantic consequences of this syntax when learning new words by at least 30 months. Mandarin children do not demonstrate the same kinds of assumptions as English children with regards to plurality and individuation.

### 13.4 Possible triggers for a parametric distinction

Having described the course of children’s acquisition of mass-count syntax and semantics in English—that children acquire the mass-count distinction between 24 and 30 months of age—we now focus our attention on possible triggers: What factors lead English and Mandarin speakers to such different syntactic and lexical systems? In this section we discuss two possibilities. The first involves the use of morphosyntactic cues. As we argue in section 13.4.1, few morpho-syntactic cues that
are available to children at the age of 2 could alone indicate to a child that their language contains count syntax. Many correlates of count syntax are also found in some classifier languages that clearly lack count syntax, such as Western Armenian.

Another possible trigger relies on children reasoning inductively about how adults use the bare noun (in English, the singular noun) to refer to singularities. The key distinction between mass and count nouns, in our view, is not that count nouns can refer to sets of countable individuals (mass nouns can too). Instead, it is that only count nouns can denote sets consisting exclusively of single individuals. By beginning acquisition with the assumption that reference to singulars should be reflected grammatically, children learning English could locate the corresponding singular-plural distinction as a first step towards acquiring the mass-count distinction. Mandarin-speaking children, in contrast, will find no corresponding grammatical correlate of singular reference, and thus no inductive basis for acquiring a singular-plural distinction or the superordinate distinction between mass and count. The details of this second possibility are outlined in section 13.4.2.

13.4.1 Problems for a morphosyntactic trigger

Although we cannot show definitively that there is no purely morpho-syntactic trigger for the acquisition of count syntax, we do believe that the most salient differences between mass-count and classifier languages are unlikely candidates. These cues include the presence of a rich classifier system, the presence or absence of a plural morpheme, and the direct combination of numerals and noun stems. As we demonstrate in this section, none of these cues is sufficient either because they are found in both types of language, or because they are not available to children at 2 years of age, when they begin to acquire count syntax in English.

13.4.1.1 Classifiers as triggers  One of the major differences between Mandarin and English is the presence of a rich classifier system. This obvious sharp contrast between these two languages is one possible way that a child might distinguish the type of language he is acquiring. However, there are two problems with this hypothesis. First, Mandarin children appear to struggle with mastering the intricacies of its classifier system until at least the age of 5 or 6. For example, Li, Barner, and Huang (2008) report that Mandarin children become sensitive to the shape specifications of classifiers only by around 3 years of age, and don’t exhibit adult-like interpretations until around the age of 6 (for related findings see Cheung, Barner, and Li 2010, Uchida and Imai 1999, Matsumoto 1985, Erbaugh 1986).

Second, classifiers share many distributional similarities with measure words in mass-count languages. On the surface, it is difficult to tell the two systems apart. Consider the measure noun phrases listed in (21).
(21)  a. Three cups of milk, one gram of saffron, two pounds of iron, etc.
      b. Two items/pieces of furniture, two items/pieces of clothing, etc.

As shown in (21a), English measure words mediate the relationship between numerals and nouns just like classifiers in Mandarin. As shown in (21b) and in section 13.2.3.2, in English there are default measure words that behave much the same way as default classifiers like ge in Mandarin. Although there are subtle differences between measure nouns and classifiers given a wider grammatical context (see Cheng and Sybesma 1999), children as old as 3 are almost completely insensitive to such differences, and the addition of a default classifier like ge has no effect on how they interpret a novel noun. In particular, children under the age of 6 do not infer that a novel noun refers to a kind of object from the presence of default classifiers (Cheung et al. 2010). Thus, it is very unlikely that children could use the presence of classifiers in their language to infer that it does not contain count syntax.

In summary, the hypothesis that a rich classifier system can serve to trigger the course of acquisition is problematic due to late acquisition of classifier systems as well as the similarities between classifiers and measure words. Any adequate theory that would want to maintain this hypothesis would have to specify the distinctive feature of classifiers that is detectable by children between the ages of 24- and 30-months.

13.4.1.2 Numeral composition and plural marking as triggers Two additional differences between Mandarin and English involve the composition of numerals and plural marking. English allows numerals to combine directly with nouns (no mediating classifier is needed) and also has a systematic plural/singular distinction. In contrast, Mandarin usually requires classifiers to mediate the combination of nouns and numerals and, although it has an infrequently used plural marker, it does not have an obligatory plural/singular distinction, as noted in section 13.2. It is possible that either (or both) of these characteristics act as a trigger in determining the type of language a child is acquiring.

However, as we look at other languages, the plausibility of this diminishes. Particularly relevant is Western Armenian (but see also Turkish as discussed in Bale, Gagnon and Khanjian 2010, and Korean as discussed in Kang 1994). Western Armenian is like Mandarin in that it has a rich classifier system that can be divided into non-sortal (22a) versus sortal (22b) subcategories. It even has a default classifier (had) similar to the Mandarin classifier ge (see 22b).

(22)  a. jergu kilo xentsor
      two cl apple
    ‘two kilograms of apples’
b. jergu had xentsor
two cl apple
‘two apples’

Like the classifier ge, *had* can be used with a variety of different nouns and does not provide any specific semantic content regarding the status of individuals in the denotation. Rather, the noun itself specifies what is to be counted, namely the atomic minimal parts in its denotation.

Western Armenian, like Mandarin, does not have any quasi-cardinal quantifiers that apply exclusively to one class of noun, nor does it have minimal noun-pairs such as *chair* vs. *furniture*, pairs which can both be used to refer to individuals but which have different syntactic patterns. These distributional facts are characteristic of classifier languages.

Also like Mandarin, Western Armenian does not make a systematic singular/plural distinction. Although Western Armenian has a plural marker, *-ner*, that is similar to the Mandarin morpheme *-men* (see the example in 23a), the so-called singular nouns in Western Armenian—nouns without plural marking—do not have a singular meaning. For example, the bare noun *dægha* (boy) is underspecified with respect to number (see Donabédian 1993, also Borer 2005). The sentence in (23b) can be used to talk about one boy running or several boys.

(23) a. Dægha-ner vaze-tsin.
boy-pl run-past(3.pl)
‘More than one boy ran.’

b. Dægha vaze-ts.
boy run-past(3.sg)
‘One (or more) boys ran.’

c. Aram-ə dægha e.
Aram-def boy is.
‘Aram is a boy.’

Aram-def and Nanor-def boy are
‘Aram and Nanor are boys.

Similarly *dægha* can be predicated of individuals as in (23c) or groups as in (23d), just like the object-mass terms such as *furniture* in English and the common noun terms such as *haizi* (child) in Mandarin.

Another similarity between Mandarin and Western Armenian concerns the behaviour of nouns in comparative constructions. As demonstrated in (24a), the noun *xentsor* (apple) is often used with the default classifier, *had*, when it is combined with a numeral.
As demonstrated in (24b), the bare noun *xentsor* can appear in comparative constructions with neither classifier syntax nor plural morphology. When it does, the comparison is made in terms of the number of individual apples rather than in terms of a continuous measurement of mass or volume, just like the object-mass term *furniture* in English or the common noun term *pingguo* (apple) in Mandarin (and unlike the noun *apple* in English).

In summary, Western Armenian has almost all of the distributional and semantic characteristics that distinguish Mandarin from mass-count languages like English. This evidence supports the hypothesis that Western Armenian does not have English-like count syntax. Rather, like in Mandarin, all the object-denoting nouns in Western Armenian have the distributional and semantic characteristics of English object-mass nouns (as discussed in section 13.2.3).

Although Western Armenian is similar to Mandarin in these respects, there are two significant differences. First, unlike the Mandarin plural marker (*-men*), which is not commonly used and can only appear on nouns with animate, human denotations, the Western Armenian plural marker (*-ner*) is frequently used and can attach to any noun that can refer to individuals (see Bale and Khanjian 2009, and Bale, Gagnon and Khanjian 2011, for a detailed discussion). Second, the combination of numerals with nouns in Western Armenian does not always require a classifier. This is demonstrated by the phrases in (25).

The use of classifiers in Western Armenian is optional. Although the direct combination of numerals with nouns is a characteristic more commonly associated with mass-count languages, there are other classifier languages that demonstrate similar properties (see Kang 1994). Despite this one difference from Mandarin, when all characteristics are taken into consideration, Western Armenian is clearly more like Mandarin than English.
Given these characteristics of Western Armenian, it seems unlikely that plural marking or the lack of classifiers in noun phrases with numerals could indicate to a child that their language makes a mass-count distinction. Like Mandarin, Western Armenian does not have count syntax. Yet the distributional characteristics of plural marking and numeral modification resemble English, at least in terms of surface distribution (although crucially not in terms of meaning). If such characteristics were to trigger count syntax, then one would expect the syntactic and semantic facts of Western Armenian to be quite different: there should be a systematic singular-plural contrast; there should be minimal pairs like chair vs. furniture; there should be quasi-cardinal determiners that apply to only one class of noun; there should be measure phrases instead of a classifier system; the lack of plural marking (or classifier marking) in comparatives should trigger coercions; plurals should be obligatory in measure phrases; etc. Western Armenian has none of these characteristics. To maintain the hypothesis that plural marking and numeral composition are triggers for count syntax, one would have to explain how a child can detect the difference between English and Western Armenian plural vs. non-plural marking, or between the two types of numeral composition. In section 13.4.2, we explore one possible way that a child might be able to do this. Critically, this possibility requires that the child use semantic cues to distinguish English plural marking from Western Armenian plural marking.

13.4.2 Induction on the use of bare nouns

We have argued that the key distinguishing property of mass-count languages is that they permit nouns which denote only sets of individuals without any groups (e.g. singulars). In classifier languages like Mandarin and Western Armenian, such meanings are not possible—any noun can be used to refer to both groups and individuals. This observation suggests a possible mechanism by which children could identify count syntax in their language. If a child were provided with evidence that certain nouns in their language input have only singular denotations, this might allow them to infer that their language makes a mass-count distinction.

Given only positive linguistic evidence, it is impossible to infer deductively that a noun’s denotation only contains single individuals. Even if it is obvious in a particular context that the noun is being used to refer to a singular thing, such a usage is also consistent with a nominal denotation that contains both groups and individuals. For example, in the utterance ‘Look at the blicket!’ the novel word blicket could be replaced with either a singular noun (e.g. cat), an object-mass noun (furniture), or a Mandarin bare noun that is consistent with any quantity of things. Assuming a deductive learning mechanism, a child who did not yet know whether her language makes a mass-count distinction would be unable to decide from a single utterance (or even from many) which grammatical representation best fits the available data.
A more plausible strategy involves inductive inference. For example, given a conspicuous use of certain nouns to denote only single individuals without a plural, a child could infer inductively that such bare nouns can only denote a set of individuals. This inference in isolation would trigger the difference between Armenian and English, and would account for the key difference between mass-count and classifier languages more generally. This type of inference, based on conspicuous correlations in language input, is ubiquitous in other areas of language acquisition. Inductive inference is necessary for acquiring category labels and extending them to novel referents (e.g. to label an unfamiliar cat as a cat). Also, it is likely needed to distinguish terms like cat from superordinate terms like animal. In one demonstration of this, Xu and Tenenbaum (2007) showed that when adults were provided with novel category labels that referred consistently to a specific type of dog, they were unlikely to infer that the label meant dog, but more likely to infer that it meant dalmatian or poodle, even though each type of thing is consistent with the dog hypothesis. As Xu and Tenenbaum show, inductive processes that capitalize on conspicuous correlations are readily implemented in inferential Bayesian models that mirror human learning.

Returning to our example, consider the data that children receive regarding the singular nouns chair and furniture. For the sake of discussion, we will limit our attention to predicative, definite and existential uses of these nouns. Given the intricacies of predication (see Gillon 1992, Scha 1981, Schwarzschild 1996, Lasersohn 1995, Langendoen 1978, among others), it would be difficult for the child to determine whether use of the noun with other quantifiers (such as the universal) would involve quantification over singular individuals or groups. The sentences in (26a–d) and (27a–d) are representative of utterances that a child might hear with chair and furniture.

(26)  
   a. Look, a chair is on the balcony.  
   b. The chair is on the balcony.  
   c. That chair is dirty.  
   d. That thing on the balcony is a chair.

(27)  
   a. Look, some furniture is on the balcony.  
   b. The furniture is on the balcony.  
   c. That furniture is for sale.  
   d. Those two things on the balcony are furniture.

Consistent with the Gricean Maxim of Quantity, the sentence in (26a) would typically be used when there is one chair on the balcony or at least when there is only one chair salient to the conversation. Also, due to the semantics of the definite determiner and the demonstrative, the sentences in (26b) and (26c) are exclusively
used to refer to a single chair. Furthermore, when the noun appears in predicate position as in (26d), it can only be truthfully applied to individuals (rather than groups).

This pattern with chair contrasts sharply with the pattern for furniture. As shown in (27), the noun furniture is often used to quantify over groups (27a), refer to groups (27b,c) and also as a predicate of groups (27d).

It is our hypothesis that children, whether acquiring a language like English or one like Western Armenian, will notice that there is a plural morpheme in their target language: a morpheme that often correlates with reference to or quantification over groups. By noticing the correlation between nouns that denote single individuals and nouns that lack plural morphology, children could inductively infer that single individuals cannot occur with plural morphology, thereby acquiring a singular-plural distinction. The steps in this acquisition process are outlined in (28).

(28) STEPS IN THE ACQUISITION OF COUNT SYNTAX
   a. The child hears plural being used to talk about groups/pluralities of objects.
   b. The child notices that whenever bare forms of nouns are used, they refer to single objects.
   c. The child infers inductively that single objects cannot be referred to by plural nouns, and therefore that plurals are used only for reference to groups.
   d. The child concludes that his target language has count syntax.

This acquisition process correctly predicts that children exposed to languages like Western Armenian or Mandarin should not acquire a mass-count distinction. First, Mandarin has no systematic plural marker (the marker men is restricted to human animate nouns). Second, as discussed in sections 13.2 and 13.4.1, nouns which can be used to talk about single individuals are often also used to talk about groups in both Western Armenian and Mandarin. Recall the examples from (11), (16) and (23) repeated below.

(29) a. Haizi pao (le).
    Child run (ASP)
    ‘The children/child ran.’

   b. Zhangsan he Lisi shi haizi.
    Zhangsan and Lisi be child
    ‘Zhangsan and Lisi are children.’

    Boy run-3-PAST(3.SG)
    ‘One (or more) boys ran.’

    Aram-DEF and Nanor-DEF boy are
    ‘Aram and Nanor are boys.’
Common nouns like Mandarin’s *haizi* ‘child’ and Western Armenian’s *dogha* ‘boy’ are often used without plural morphology to quantify over and refer to groups ((29a) and (30a)) and are also often used as a predicate of groups ((29b) and (30b)).

In summary, the acquisition procedure outlined in (28) could be used to reliably trigger the acquisition of count syntax for English children while also correctly not triggering such an acquisition for Mandarin and Western Armenian children.

### 13.5 Conclusion

As argued in section 13.2, the mass-count distinction in English does not reflect a simple ontological distinction between denotations with and without individuals. Rather it is a syntactic distinction that has subtle semantic implications. Critically, individuation is not exclusively encoded by count syntax. Mass syntax permits individuated denotations as well (*furniture, equipment*, etc.). Yet, there is an essential difference between individuated denotations associated with mass syntax and those associated with count syntax: singular nouns with count syntax only denote a set of single individuals whereas those with mass syntax denote both individuals and groups.

As argued in section 13.4, the presence of plural morphology alone does not reliably signal count syntax. However, languages with count syntax always have both plural morphology and singular nouns that denote sets of individuals. In acquisition, children could infer that their language contains count syntax by observing that it makes a distinction between plural and non-plural nouns, and that some of the non-plurals denote singleton sets. This inference would license a host of additional inferences about the language, none of which appear to drive children’s initial acquisition of the distinction.

If the language has count syntax, then the child will know to be sensitive to selectional distinctions based on this syntax (*each vs. all*). They will know that some variations in form depend on this syntax (*much vs. many*), that there is no classifier system but rather a variety of measure nouns, that there are two types of singulars (count vs. mass singulars), and that there are two types of nouns that encode individuation (object mass-nouns, regular count nouns), not to mention all the syntactic and semantic consequences that follow from this knowledge.

In contrast, a child who fails to detect count syntax has inferred that their language contains only one type of noun that encodes individuation, and that all determiners in the language select for nouns based only on semantic distinctions (not syntactic distinctions). The child will also know to be sensitive to the presence of a rich classifier system that could possibly be divided syntactically as well as semantically into two categories: sortal vs. mensural.
Furthermore, the lexical acquisition strategy children employ is in part determined by the properties of the language they are exposed to. Children that fail to detect count syntax have a straightforward way of gathering grammatical information about a word from how it is used. For words that are used to refer to individuals, children can infer that they are nouns with an individuated denotation. For words that are not used to refer to individuals, children can infer that they are nouns without an individuated denotation. Such information aids the child in acquiring a lexicon.

In contrast, children that have acquired count syntax do not have such a straightforward way of gathering grammatical information. For words that are used to refer to non-individuals, children can infer that they are mass nouns with non-individuated denotations. However, for words that are used to refer to individuals, children have two possible options to consider: such words might be mass nouns with individuated denotations or they might be count nouns with singular denotations. Children would need to pay attention to other aspects of how the word is used (e.g. use in singular syntax to denote only singularities) to determine which of the two is more likely.
Classifying and massifying incrementally in Chinese language comprehension*

NATALIE M. KLEIN, GREG N. CARLSON, RENJIE LI, T. FLORIAN JAEGER, AND MICHAEL K. TANENHAUS

14.1 Background

Both mass and count nouns in Mandarin require a classifier when following a numeral, a universal quantifier, or a demonstrative, as in (1) and (2):

(1) Na ben shu hen gui.
    that CL-volume book very expensive
    ‘That book is expensive.’

(2) Wo he le liang wan tang.
    I drink past two CL-bowl soup
    ‘I ate two bowls of soup.’

Which particular classifier is used for any given noun is grammatically rather than semantically determined, although the classifier may have an inherent meaning of its own and may evoke salient properties of the nouns with which it combines (Huang and Ahrens 2003). Scholars have categorized classifiers and classifier/noun pairings in several ways. Huang and Ahrens (2003) distinguish classifiers by whether they coerce individual, kind, or event meanings of the noun. Gao and Malt (2009) group classifiers by how evocative they are of the conceptual features of the nouns with which they pair. For example, they identify classifiers that evoke well-defined properties, prototype classifiers, and arbitrary classifier groups. Gao and Malt’s (2009) work raises the important observation that Chinese classifiers are unlikely

* This research was partially supported by NIH grant HD27206. Correspondence to nklein@bcs.rochester.edu. We thank Roger Levy for discussions that highlighted the implications of differences in the informativity of generic and specific classifiers.
to be uniformly informative about which noun might follow, with some being best viewed as generic or default (Myers 2000).

In this chapter, we will primarily focus on Cheng and Sybesma’s (1998) taxonomy, which sorts classifiers based on ontology (see Krifka 1995, and Yang 2001, for similar divisions). The authors distinguish two main types of classifier use in Chinese: count classifiers, which combine with nouns to denote entities that appear individuated in the world, and mass classifiers or ‘massifiers’, which combine to denote substances or unindividuated groups. While the count classifiers have been of more interest to psychologists because of their potential relationship with categorization and conceptualization, massifiers are more analogous to English measure phrases, such as ‘three bouquets of lilies’ or ‘that bottle of wine.’ Although, unlike Gao and Malt (2009), Cheng and Sybesma’s (1998) division only distinguishes two types of classifier and ignores salient conceptual dimensions, it is not a cognitively arbitrary division: the authors assume a pre-linguistic ontological distinction between mass and count concepts.

Yet the mass/count distinction is not syntactically realized in Mandarin and all but a few exceptional nouns require classifiers following numerals or determiners. Given this, Chierchia (1998b) and others have suggested that perhaps all nouns in Chinese are basically mass. Even if this is so, it does not follow that Chinese nouns cannot be conceptually distinguished as massifying and classifying. It is possible that a conceptual mass/count distinction may manifest itself as a difference in the way the noun meanings are understood by the comprehender. In this chapter, we take an experimental approach to the following question: is there evidence that classifiers and massifiers exert a different influence on language processing? And if so, what is the nature of the difference? Since there is little experimental work on classifier processing for us to rely on, we instead look to research on related structures to give us an idea of what we may find. Grammatical gender is one such area we might consider. Classifiers resemble grammatically gendered determiners in notable ways: they are required before nouns in certain syntactic environments, and they sometimes align with natural categories but are often arbitrary (Vigliocco et al. 2005, Zubin and Kopcke 1981). Evidence has shown that gender-marked determiners constrain possible referents during processing. Dahan et al. (2000) demonstrate such an effect in a series of experiments employing a visual world task that investigates gender-marked determiners in French. The presence of a gender-marked determiner in these studies mediates the strength of competition from phonological cohorts, but does not prompt early anticipatory looks to gender-consistent objects in the display. Chambers et al. (2002), who primarily examine whether physical properties of objects can impact referential selection online, also demonstrate that definiteness and article form constrain utterance interpretation, but do not necessarily increase predictions listeners make during comprehension. These and other studies suggest that articles, and perhaps classifiers and other grammatical function words by
extension, guide interpretation only to a subtle, context-dependent extent. (Although cf. Lew-Williams and Fernald (2007) for a case where grammatical gender plays a potentially larger role in children’s Spanish comprehension.)

However, the parallel between articles and classifiers is not complete: for example, determiners and articles are by nature short words that are often phonetically reduced in speech. While Chinese classifiers are usually monosyllabic, they also carry a lexical tone and are therefore likely to be longer and more acoustically prominent. Moreover, the small set of (even gender-marked) articles makes the statistical informativity of such words extremely low. For example, the number of words able to follow the, or even le or la in French, is high. By contrast, there are many more Chinese classifiers in the lexicon, and some of them only commonly pair with two or three nouns. This increase in statistical informativity may allow classifiers to play a larger role in comprehension than determiners. Classifiers also intuitively bear content and meaning in many circumstances, so in the end we would expect classifiers to have at least as much influence on sentence comprehension (and possibly more) as gender-marked articles. The degree of this influence is also potentially different for English measure phrases on the one hand, which are typically longer and involve words that readily appear as both nouns and measure phrases (e.g. ‘box’) and are therefore temporarily of ambiguous function, and Chinese massifiers and count classifiers on the other, which appear in positions that nouns cannot. The following studies we report investigate the time-course and nature of classifier and massifier comprehension, and focus in particular on four questions: To what extent do English measure phrases influence language comprehension online? To what extent do classifiers in Mandarin Chinese mirror these effects? Do count and mass classifiers have an equivalent impact on language processing? Do these effects occur at a structurally superficial level, or do they reflect deeper syntactic and semantic processes?

14.2 Previous research

Recently, some researchers have begun to examine the comprehension of mass and count classifiers in Chinese, although this is still quite novel. Li, Barner, and Huang (2008) investigate the ontological distinction in an offline judgement task and conclude that mass/count properties take precedence over other features, such as shape, in overt decisions about classifier-noun pairing. Subjects were shown a quantity of substance formed from a clay-like material into a good shape match for a particular count classifier while being an obviously malleable mass. When told that a character wanted ‘one CL-shape something’, adults opted to choose a closed mystery box as the best match for the shape classifier rather than choosing the mass substance formed to be an appropriate shape match. The authors conclude that comprehenders’ preference for the mystery box stems from a salient mass/count distinction that affects the conceptual goodness-of-fit of classifiers to nouns.
However, this task allows for overt contemplation during a forced choice with novel objects. It is possible that these circumstances call certain conceptual properties to the mental forefront. Subjects are in a position where the mystery box is salient in the environment, and where the shape-appropriate substance is not a known or named object associated with any particular classifier. In cases like this, comprehenders may rely on the mass/count distinction over shape information, but this does not necessitate that a mass/count distinction impacts language comprehension in an immediate fashion. Other groups have examined specifically the online effects of processing classifiers, although none has previously addressed referent ontology. Huettig, Bowerman, and Majid (2008) used a visual-world task to investigate a strong version of the linguistic relativity hypothesis, and found that, when a named target item is not present in the display, comprehenders look more readily at classifier-consistent items, but only when the classifier is present in the utterance. Since this classifier-consistent item—which in Huettig et al.’s (2008) study was intentionally selected to have minimal featural overlap with the named target—did not gain increased looks when the classifier was not explicitly mentioned, the authors argue against strong linguistic relativity. The authors conclude that, while they find no evidence for strong linguistic relativity, comprehenders do use classifier information to predict upcoming referents.

While this work does seem to show some effect of classifiers during online processing, this study cannot exclude the possibility that subjects merely looked at items that were the most phonologically consistent with the input. At the surface level, the sounds present in the classifier might map best to the classifier-consistent picture when no other picture in the display is a phonological or conceptual match for the instruction. These data do not rule out the possibility that classifiers are being treated as a clitic or prefix, or that the subjects’ eye-movements are reflecting phonotactic probabilities at a shallow level. In other words, Huettig et al.’s study does not tell us that comprehenders process classifiers in a grammatically relevant way, or that semantic information from classifiers matters online. Tsang and Chambers (2011) have used eye-tracking to examine the role of Cantonese classifiers in sentence comprehension. Cantonese classifiers have somewhat different discourse properties than Mandarin classifiers (Cheng and Sybesma 2005), but have strong lexical similarities and nearly identical noun groups. Tsang and Chambers (2011) primarily investigate the extent to which classifiers evoke shape properties online by using a visual-world task with shape-similar distractor objects. One interesting note these authors make is that the link between classifiers and shape seems surprisingly tenuous in their studies. Much of the literature on classifiers cites shape as a dominant conceptual correlate of Chinese classifiers, and Tsang and Chambers’s (2011) results call into question how strongly this influences the earliest moments of processing. Although many possible explanations exist for a null result, this same paradigm has demonstrated the influence of shape-competitors on word recognition
(Dahan and Tanenhaus 2005), so it seems unlikely that the design of the authors’ studies is simply not sensitive enough to capture such effects with classifiers. The unexpectedly weak relationship between classifiers and shape in these studies suggests a more complex relationship between classifiers and conceptual properties, and classifiers and corresponding nouns. Some authors, including Gao and Malt (2009), have explicated a structured hierarchy of classifiers based on dimensionality of the physical features they evoke, as opposed to mere shape, and it is possible that Tsang and Chambers’s (2011) methods would reveal differential effects based on this taxonomy. Put another way, the visual properties of an object that a classifier evokes are not necessarily predictable from an intuitive examination of the object, and so may not be the most important or reliable feature of the object in all contexts; thus while there is an intuitive relationship between classifiers and object shape, this link may not be especially informative to the comprehender as an utterance unfolds. Unlike count classifiers, which often pick out one of many object properties, massifiers provide mensural circumscription for collective referents or unindividuated substances (Tai 1994). Rather than highlighting a particular property internal to the referent, massifiers are indicative of the quantity, configuration, or external containment of the referent—and these properties can be temporary. It is possible that, because true mass referents (such as substances) must be semantically partitioned or linguistically contained to be denoted, measure phrases and massifiers are processed differently than count classifiers.

14.3 Experiment One: English measure phrases

We initially sought to gain experimental traction in this domain by examining the online comprehension of English measure phrases. English has a grammatical mass/count distinction that leads to the use of measure phrases to denote mass referents. These measure phrases, as in (3) and (4), have a word order that is similar to equivalent Chinese phrases. Note that (3) refers to a grouped quantity of count referents, while (4) refers to a true substance.

(3) Harriet smoked two packs of cigarettes.

(4) Gretchen drank a bottle of gin.

These measure words come from an open class and can refer explicitly to a container item or a group array when examined as isolated words. Because of the strong intuitions that English measure phrases are semantically contentful, we might expect a prominent effect of the measure word on referent resolution, where listeners use the measure phrase to begin to eliminate incompatible referents and/or make guesses about compatible referents before hearing the noun itself. In order to gain fine-grained temporal information on the comprehension of these measure phrases, we use the visual-world paradigm, which allows us to infer participants’
understanding of our linguistic materials at any point in the speech stream by gauging where the subject is looking (Tanenhaus et al. 1995).

**Participants:** All 16 participants were native speakers of English from the University of Rochester community. They had normal or corrected-to-normal vision and reported normal hearing, and were paid for their participation in the experiment.

**Materials:** The experimental task involved selecting a target from a display of four groups of pictures on a computer screen, as shown in Figure 14.1. A target group, its phonological competitor group, its classifier competitor group, and the classifier competitor’s phonological competitor group were always present in quadrants on the monitor. Fourteen critical instructions containing a measure word, such as ‘Choose a head of lettuce’, were given to subjects aurally along with a unique array of pictures. The targets were selected for imageability and familiarity, and were either substance or collective items. A second item that pairs with the same linguistic measure word was chosen as the classifier-consistent competitor. For example, the measure phrase ‘a head of’ was paired with ‘lettuce’ as a target and ‘broccoli’ as a classifier competitor. For each of these nouns, an imageable phonological cohort was chosen, which shared initial phonology—but could not typically be paired with—the

---

**Figure 14.1** Screen shot of English measure ‘head of’ referential selection trial, which shows four groups of pictures: lettuce, letters, broccoli, and broth, and a central fixation cross. Instructions in the general condition were ‘Choose a picture of lettuce’. Instructions in the specific condition were ‘Choose a head of lettuce’.
target measure phrase. ‘Letters’ was matched with ‘lettuce’ and ‘broth’ was matched with ‘broccoli’, for instance.

Each group of pictures had three identical tokens of an object, in order to maximize the felicity of using a measure phrase for reference. That is, there were three heads of lettuce on the screen so that an instruction to ‘choose a head of lettuce’ is neither an awkward use of the indefinite nor emphatically contrastive with the other items on the screen. Trial order and picture quadrant were presented in a randomized order and at randomized locations on the screen for each participant, respectively. A fixation cross was located in the middle of the screen to allow the eye-tracking software to recalibrate periodically throughout the experiment. Instructions were recorded by a naive native speaker at two different levels of classifier specificity. A general, or uninformative-classifier instruction, such as ‘Choose a PICTURE of lettuce’ was matched with a specific, or informative-classifier instruction like ‘Choose a HEAD of lettuce’. The only difference between the general and specific condition was the degree of informativity of the measure phrase: in the general condition, all pictures in the display were consistent with the measure phrase (e.g. ‘picture’), and in the specific condition, only the target and the classifier competitor were consistent with the measure phrase (e.g. ‘head’ or ‘bottle’). Half of the critical items subjects received had a general classifier and half had a specific classifier. Fillers, which were randomly mixed with critical trials, referred to one of the phonological cohorts as a target, and were equally likely to be presented with a general as with a specific instruction. This was to prevent subjects from being able to identify the target picture as being one that had both a phonetic and a classifier-consistent competitor on the screen concurrently.

**Procedure:** After giving informed consent, participants were fitted with a head-mounted eye-tracking device and headphones. After calibration to the subjects’ eyes, participants were told that they would hear instructions to click on certain pictures they would be shown. Three practice trials appeared, and unless the subject had any questions at that time, the experiment began immediately. Pre-recorded stimuli played over the headphones at a comfortable volume, and there was a short break halfway through the study to prevent ocular fatigue. The study took approximately thirty minutes, after which participants were debriefed and paid for participation.

**Predictions:** This study compares two levels of measure-word informativity, general and specific. In the general condition, all items in the display are consistent with the instruction up through the measure phrase, which provides no real information about which picture is the target. Comprehenders must wait for the noun to begin to select a target. In this condition, we expect to see standard phonological cohort effects, where participants will begin to look at the two similar-sounding groups shortly after noun onset. The target ‘lettuce’ and its phonological cohort ‘letters’ should compete for selection until a phonetic point of disambiguation (POD) can be
processed, approximately 200 milliseconds after it appears in the auditory stimulus, which is the approximate time necessary to plan and launch a saccade.

We expect participants to behave differently in the specific instruction condition. Because 'head of' is only consistent with 'lettuce' and 'broccoli', looks to those two items might rise soon after the classifier onset. Because the measure phrase is not consistent with the phonological cohort, 'letters', the usual cohort effect should be absent in this condition, if participants rapidly integrate measure-word information when understanding reference to mass items. If, however, measure words do not have a rapid impact on language comprehension, we may still see signs of phonological competition in the specific condition, or we may just not detect any increased attention toward the classifier-consistent competitors like 'broccoli'.

Results: To examine the effects of measure phrases on sentence comprehension, we aligned all trials to the onset of the noun and then plotted the proportion of looks to the four object groups on the screen as the instruction is uttered. As predicted, we found evidence of phonological competition in the general condition, as shown by increased looks to both the target and its phonological cohort approximately 200 milliseconds after onset of the noun. While looks to the classifier competitor and its cohort pair begin to fall at this point, looks to the phonological competitor distinctly do not start falling until around 500 ms after noun onset, in other words, not until the phonological point of disambiguation is processed. Proportions of looks to the four item groups in the generic condition are shown in Figure 14.2.

![Figure 14.2](image-url)
In the context of the more restrictive specific measure words, the pattern of looks is different. Just before the onset of the target noun, looks to the two classifier-consistent groups already attract approximately 70 percent of fixations. No noun has been mentioned, but information present in the measure word has impacted subjects’ expectations about what will be referred to in the instruction. This is shown in Figure 14.3. The target diverges from the classifier competitor around 275 ms after noun onset, which is faster than in the general condition, and shortly after noun onset but before the phonological POD. Looks to target groups in the two conditions are plotted together in Figure 14.4.

**Discussion:** The pattern of looks at the noun onset in the general condition corroborate the idea behind our stimuli: participants are looking at all four items about equally often. Our generic measure phrase was truly uninformative, as intended, and looks begin to diverge approximately 200 ms after the noun begins, which is very much in line with what we know about the time-course of planning and executing saccadic eye-movements. After this point, looks to the target rise, looks to the phonological cohort somewhat increase, and looks to the other items drop off, because they are incompatible with the first sounds of the noun. Eventually, as a word like ‘lettuce’ disambiguates from cohort ‘letters’, looks to everything but the target drop away. This describes the proportion of looks at about 500 ms after noun onset, and reflects phonological cohort effects widely found in the word-recognition literature.

![Figure 14.3](image_url)

**Figure 14.3** Proportion of all looks to the four picture types over time (ms) in the specific condition, aligned to noun onset
By comparison, participants have already begun to use information present in the specific measure phrase before they hear the noun at all. Looks to the target begin to rise almost immediately after noun onset, which is faster than it would take to plan a saccade in response to information in the noun; these looks are rising as a response to the measure phrase. By 300 milliseconds post onset, the target group is receiving the majority of looks. Interestingly, at no point does the phonological cohort show a strong advantage over the fourth, unrelated item. Even though the initial phonology of the target noun is consistent with that cohort, it is not a competitor for selection in the presence of a classifier that restricts it. Since the phonological cohort is ruled out as a competitor by the measure phrase, the pragmatic point of disambiguation is earlier than the phonological POD. This manifests as quicker target selection in the specific condition, as shown in Figure 14.4.

This study provides evidence that, in English, measure phrases that semantically partition collectives or substances for reference have an influence on the earliest moments of sentence comprehension. These effects mediate phonological cohort competition when the measure phrase is inconsistent with or atypical of the cohort. But there is reason for caution when interpreting these results: in English, measure

![Graph showing English Mass Classifier Target Looks](image)

**Figure 14.4** Proportion of all looks to the target groups over time (ms) in the two conditions, aligned to noun onset

---

1 It is also notable that in Figure 13.3 the classifier competitors are more attractive initially than the target pictures. In order to avoid giving target selection a confounding advantage, when in doubt, we erred on the side of assigning the better classifier-fitting object as the competitor and not the target. Therefore the initial advantage of the classifier competitor relative to the target in the specific condition is likely to be a product of our items.
words often resemble nouns. In the example we have been discussing, it might seem odd to refer to a head of lettuce as simply ‘a head’, because there is a more dominant sense of the noun ‘head’. But for many other measure words in our study, including ‘carton’, ‘bottle’, and ‘cup’, the target picture could, if in isolation, be named felicitously with this measure term alone. Because of the temporal nature of speech, the syntax of a measure phrase and that of a mere container noun are temporarily ambiguous in English. While prosodic factors and experience with earlier trials make it unlikely that subjects believe ‘cup’ is the end of an intonational phrase, we cannot discount the role this ambiguity might play in the strength of these effects.

Chinese shares this problem of lexical ambiguity: many (especially mass) classifiers are identical in form, and clearly related, to nouns. However, Chinese syntax disambiguates a classifier phrase from a noun phrase at an earlier point in time: only a classifier, and not a noun, can directly follow a numeral. This property will allow us to test more clearly not only the influence of classifiers on language processing, but the structural level at which this takes place. Chinese massifiers may behave very much like English measure phrases, due to reference to external configuration properties; or they may behave more like grammatical gender determiners, because they are function words required by the grammar for reference to a nominal concept.

14.4 Experiment Two: Chinese massifiers

The second study replicates the methods of the first study with Chinese mass classifiers, or massifiers. Many of these massifiers are analogous to English measure phrases. One crucial difference is that, unlike the nouns that can share their form, massifiers can occur immediately following a numeral or demonstrative term. Thus the presence of a numeral grammatically dictates that the following word cannot be a noun. Whether this cue is used by comprehenders during processing has never been investigated, so Experiment Two allows us to probe measure phrase comprehension at the level of syntactic processing. As mentioned, some previous research has examined how Chinese mass classifiers are paired with potential referents (e.g. Li et al. 2008). Those data suggest that comprehenders make use of ontological information in offline judgements about massifier-noun pairs. This experiment investigates whether grammatical massifiers impact comprehension online as an utterance unfolds.

Participants: All 18 subjects in this study were native speakers of Mandarin Chinese from the University of Rochester community who had normal or corrected-to-normal vision and no reported hearing loss. All were also minimally proficient in English, and many spoke a second Chinese dialect. Some of these participants also participated in Experiment Three with count classifiers. Subjects received payment for participation.
Materials: Items in Experiment Two closely resembled the stimuli from Experiment One. An imageable mass or massified target was chosen (e.g. baihehua, 'lily(s)') along with a classifier-consistent competitor (e.g. meigui, ‘rose(s)’). A phonological competitor that shared the first syllable as the target was then selected (e.g. baicai, ‘cabbage’). Whenever possible, the cohort also shared the initial tone, but this was not the case for every item. Crucially, the phonological competitor did not typically pair with the target massifier.

Objects were displayed in much the same way as in Experiment One, as shown in Figure 14.5.

One important change from Experiment One is the way the fourth picture was chosen. Instead of a phonological cohort of the classifier competitor, a cohort of the massifier term itself was chosen. For example, the classifier SHU ‘bouquet’ is appropriate for the pictures we selected of roses and lilies, and so the fourth picture group is shu, ‘tree(s)’, which is a phonological match for the measure word. The purpose of this was to explicitly test whether the ambiguity between noun and classifier form can account for any rapid looks to the target, even when a noun should be ruled out by the syntactic context. If so, we might also expect increased looks to this fourth item, the classifier-cohort, in the specific condition. Hearing classifier SHU might drive attention toward the homophonous noun shu, even though a noun is not licensed by the syntax at that point in the utterance. If we found this to be the case, this would reveal that hearing a form phonologically consistent with a picture is enough for participants to look at that item; in other words, this would lend support to a superficial explanation of classifier interpretation. If, however, subjects do not look at the picture of the trees when hearing the classifier SHU, this could provide evidence that classifiers are being treated in a

Figure 14.5 Screen shot of Chinese mass classifier SHU referential selection trial, which shows four groups of pictures: target baihehua, phonological cohort baicai, classifier competitor meigui, and classifier-cohort shu, and a central fixation cross.
syntactically appropriate manner from the earliest moments in processing. In this study we made half of the target items true mass substances and the other half massified groups of discrete objects. (See Cheng this volume, for information on classifiers and different classes of noun.) Because Chinese does not syntactically distinguish mass and count nouns, we did this based on perceptual intuitions as well as by analogy to items in the preceding English study. A naive female native Mandarin speaker was chosen to act as the voice for our experimental instructions. We recorded a general instruction, such as ‘Xuan yi XIE baihehua’, which is translated roughly as ‘choose some lilies’ or ‘choose a few lilies,’ and contrasted that with a more specific instruction, ‘Xuan yi SHU baihehua’, literally ‘choose one BOUQUET lily(s)’, or ‘choose a bouquet of lilies’.

Fourteen fillers were included that appeared identical to critical items, except that the target was one of the other three item types in the display. Half the fillers used the general classifier XIE in the instruction, and the other half used a specific massifier appropriate for the referent. As in Experiment One, this was done to prevent subjects from learning that the target was a picture that had both a classifier and a phonological competitor co-present in the display.

**Procedure:** The procedure was identical to that in Experiment One.

**Results:** The data patterns in this study were similar to the English massifier data. Looks to target items do not increase until 550 milliseconds after noun onset in the general condition (shown in Figure 14.6), whereas they rise prominently within 150 milliseconds of noun onset in the specific condition (shown in Figure 14.7). Phonological cohort competition, while never especially prominent, is nearly twice as frequent in the general condition as in the specific condition. The classifier-cohort items received very few looks overall.

**Discussion:** Effects of the massifier on eye-movements occurred very early, before information contained in the noun itself could direct a saccade. This contrast between the general and specific conditions demonstrates that comprehenders can use information from mass classifiers rapidly as they are processing a sentence. These are analogous to the effects seen with English measure phrases in Experiment One, although the Chinese massifier data are somewhat more subtle.

The attention directed to phonological competitors was diminished in the specific condition, where the massifier was inconsistent with the cohort. This suggests that semantic restrictions in the massifier are excluding the cohorts as potential referents, even when phonological information in the noun phrase is consistent with those pictures. Most noticeably, we find a target advantage in the specific condition, where the information in the massifier speeds referential selection. As previously mentioned, these effects alone do not speak to the level of processing that is taking place.
To better understand the level of processing at work, we included the classifier-cohort objects, like *shu*, ‘tree’, which share phonology (although not necessarily orthography) with the specific classifier used, which is SHU ‘bouquet’ in this example. If looks had increased to the classifier-cohort in the specific condition, this could have indicated that comprehenders are using massifier information at a superficial level, where the phonological consistency of a small window of input directs attention to matching pictures. If this had occurred, part of any effects we found might be due to a boost when the classifier shares its form with a related noun that could also describe the target picture. An example of this in English would be that ‘bottle’ could felicitously refer to a picture of wine, if the wine is depicted as being in a bottle.

However, looks to the classifier-cohort were always minimal. They were, in fact, lower in the specific case, where the matching phonology was present, than in the general condition, where XIE was always used as the classifier. This is the pattern of results we expect if the presence of *yi*, the numeral, provides a clear signal to comprehenders that the next thing they hear cannot be a noun, but is likely to be a classifier. In other words, these data indicate that subjects interpreted the classifier terms in a structurally relevant way, and not as superficial phonotactic signals, nor as potential nouns. These data show that massifier affordances rapidly influence a comprehender’s set of referential alternatives with mass referents. These effects are operating at a syntactic level. The results of Experiment Two highlight the similarity of Chinese massifiers to English measure phrases: both seem to partition a set portion of a substance, or name a collective. The comprehender is able to use the information in the measure phrase very quickly to select a target before its phono-
logical point of disambiguation. Nouns that sound the same but are partitioned differently do not compete during referential selection. However, we know intuitively that this type of semantic circumscription is more necessary when referring to mass or massed referents than count referents. Thus count classifiers might not provide information that is important during sentence processing. On the other hand, count classifiers are required by the grammar of Chinese and often evoke conceptual features of the referent. Experiment Three directly assesses the influence of count classifiers on Chinese sentence comprehension.

14.5 Experiment Three: Chinese count classifiers

Experiment Three directly investigates the influence count classifiers have on Chinese sentence processing when potential referents appear as individuals, not groups or substances. Huettig et al. (2008) and Tsang and Chambers (2011) both use eye-tracking to study Chinese count classifiers. Yet both groups are concerned with absolute looks to one particular item in the display. Neither group explicitly looks at how the level of classifier informativity can impact the dynamics of competition between objects during referential selection.

Participants: Eighteen native Mandarin speakers from the University of Rochester community participated in this study for pay. All had normal or corrected-to-normal vision and no reported hearing problems. All were at least minimally proficient in English, and many spoke other dialects of Chinese in addition to Mandarin.
Materials: Stimuli in this study closely resembled those of Experiment Two with one major difference: all items displayed were conceptual count objects, not mass substances, and none of the objects appeared in cohesive groups (i.e. no ‘bouquets’ of flowers, ‘packs’ of cigarettes, etc.).

Target items, such as men ‘door’, were paired with words that take the same classifier, like chuanghu ‘window’. Phonological cohorts (menpiao, ‘ticket’) shared an onset syllable (and, as often as possible, tone) with the target but did not pair with the target classifier. The fourth item again was a phonological cohort of the specific classifier itself. In this item, men and chuanghu both pair with the classifier SHAN, which connotes flat vertical things, so the classifier-cohort was shanzi, ‘fan’. To preserve felicity and to maintain the same level of visual complexity as in Experiment Two, the individual objects were depicted in groups of three identical tokens, as in the first two experiments. This example is shown in Figure 14.8.

Again we compared a specific instruction, such as ‘Xuan yi SHAN men’, ‘Choose one CL-flat door’, with a less informative instruction. Chinese has a general classifier for individuals, GE, which we used in lieu of the restrictive classifier in our general instructions, as in ‘Xuan yi GE men’, ‘Choose one INDIVIDUAL door’. GE has been argued to be a true default classifier (Myers 2000), although acceptance of GE with a particular noun may vary, especially in formal speech or writing. More importantly, GE seems the best uninformative pragmatic equivalent for the selection of individuals.

A female native speaker of standard-sounding Mandarin recorded all instructions. Fourteen fillers, which were randomly mixed with critical trials, were of the same type as in the previous study.

Procedure: The procedure was identical to that of Experiment Two.

Figure 14.8 Screen shot of Chinese count classifier SHAN referential selection trial, which shows four groups of pictures: target men, phonological cohort menpiao, classifier competitor chuanghu, classifier-cohort shanzi, and a central fixation cross.
Results: As in Experiment Two, we were concerned with looks to each item type as information from the speech stream is processed by the comprehender. Figure 14.9 and 14.10 show the proportion of looks to each item type as the utterance unfolds.

In the general condition, when the default classifier GE was a good match for all the items shown, we see looks to all four object groups attract approximately equivalent looks until around 200 ms after the noun onset, at which point the classifier competitor and classifier-cohort looks begin to drop. The phonological cohort looks remain steady and target looks rise until around 400 ms post noun onset, at which point the phonological competitor looks begin to recede. Target items attract half of all looks by about 550 milliseconds after the noun begins. The specific classifier condition looks different. Initially looks to all four object groups are equivalent, but around 100 milliseconds after noun onset, looks to the phonological competitors begin to fall. The classifier-consistent competitor looks peak near 30 percent at approximately 350 milliseconds after noun onset. By 250 milliseconds after noun onset, the target group is attracting nearly one third of looks, and by 450 ms post onset that becomes half of all looks.

Discussion: The basic pattern of these results resembles the massifier data: phonological cohort and classifier competitor trade off as the dominant challenger for attention depending on the level of classifier restrictiveness. The target receives the majority of subjects’ looks by about 100 milliseconds quicker in the specific classifier condition. This advantage demonstrates that comprehenders make swift use of

\[
\text{Proportion of Looks}
\]

\[0 \quad 0.2 \quad 0.4 \quad 0.6 \quad 0.8 \quad 1\]

\[-50 \quad 150 \quad 350 \quad 550 \quad 750 \quad 950\]

**Figure 14.9** Proportion of looks to the four picture types over time(ms) in the general condition, aligned to noun onset
nominal classifiers during reference resolution. As with massifiers, the classifier-
cohort items earn very little attention from subjects, suggesting that facilitation
effects from the specific classifiers cannot be attributed simply to phonetic atten-
tion-driving. Notably, competition from the phonological cohort practically disap-
pears in the specific condition. This is consistent with Dahan et al.’s (2000) research
on grammatically gendered determiners. Classifier competitor looks seemingly drop
away later and to a lesser degree with count classifiers than with massifiers. Yet this
pattern is still important and indicates that, firstly, classifier specificity facilitates
comprehension, and secondly, count classifiers play a role in referential selection by
changing the set of competitors that are entertained during comprehension.

14.6 Comparing across language and ontology

Because the procedures used in the three studies were nearly identical, differences in
the data provide insight into classification across languages and ontologies. What we
see in broad strokes is that classifiers across languages and types provide contextual
information that is used in the earliest moments of sentence comprehension, although
the impact and nature of these effects varies somewhat by language and ontology. With
English measure phrases, we see a clear early influence of measure words on anticipa-
tory looks toward classifier-consistent objects. In the specific condition, just before the
onset of the noun, only about 20 percent of looks are to objects inconsistent with the
measure phrase. Notably, the target’s phonological cohort is only a competitor when
the measure word is generic enough to allow it to be a plausible target. This is unsurprising given our intuitions that English measure words are contentful. This effect is likely strengthened by the ambiguity between measure words and related nouns denoting containers in English.

Chinese massifiers play a similar role in dividing up the world for counting, but they appear in grammatical structures that identify them unambiguously as classifiers, not nouns. While some massifiers share morphology or form with nouns, these nouns cannot directly follow a numeral term, but would need their own classifier. The absence of looks to the classifier-cohort pictures indicates that comprehenders are aware of the syntactic category of the massifier and are using this information during comprehension. Therefore, early looks to the target and classifier competitors in the specific condition are not attributable to the class ambiguity of some massifiers, or to superficial phonological attention-driving.

Although Chinese massifiers have meanings that seem to map well to English measure phrases, their impact on comprehenders’ visual attention is not quite as strong. In the specific condition, where the massifier is informative enough to halve the number of potential targets, looks to all four object types are approximately the same until shortly after noun onset. This represents a delay compared with English measure phrases, although these effects still occur before information in the noun could be used to plan and launch a saccadic eye-movement. And as in Experiment One, the strength of the phonological cohort competition varies based on specificity. Massifier information in Chinese constrains comprehension quite rapidly, both by facilitating target selection as shown in Figure 14.11, but also by influencing which other items in the display compete with the target. The delay of these effects relative to English, however, raises the possibility that massifiers in Chinese direct comprehenders’ attention to consistent objects somewhat less powerfully.

Count classifiers show even more subtle effects in the same direction. The target has very little advantage in the presence of a more informative classifier, as shown in Figure 14.12. However, the type of object that competes the most strongly for selection varies with classifier specificity. Count classifiers, then, seem to offer contextual information that constrains comprehenders’ interpretations. However, the advantage in terms of target selection is minimal. Although count classifier/noun relations are more limited by Chinese grammar, and might therefore provide clearer information about what nouns can come next in the utterance, our data show that they offer less facilitation for target selection than massifiers or English measure phrases do.

We must also examine alternative explanations for the differences we see in our data other than an ontological distinction encoded by the classifier. One possibility is that the lexical co-occurrence of our Chinese mass nouns with our massifiers was
higher than the co-occurrence of our count classifier/noun pairs. If the probability of a certain noun given a classifier was higher overall for the mass nouns in our study, this could result in the faster and larger influence on comprehension of massifiers relative to classifiers.

We preliminarily investigated this by searching for classifier and classifier/noun bigram frequencies on Google’s Chinese language pages. We then divided the classifier frequency by the bigram frequency to quantify the proportion of times the noun appears when a particular classifier is given. A t-test between the log values of these proportions revealed no significant differences between the predictiveness of our massifiers and count classifiers. If anything, the count classifiers are more predictive of our count referents than our massifiers are of our mass referents by this metric \( t = -1.22, p = .232 \), but this is not statistically significant.\(^2\) There are additional reasons besides these co-occurrence data to believe that count classifiers may be generally more predictive than massifiers. Massifiers can be paired with many nouns and still maintain grammaticality and interpretability. Meanings of even odd massifier/noun pairs can be coerced, and new containers or quantities can be created and turned into massifier terms. Count classifier/noun pairings, however, are more grammatically rigid, and the use of a count classifier with a mismatched

---

\(^2\) Despite this preliminary finding, we do not want to completely rule out the possibility that, given a large spoken corpus of Chinese and further analyses, classifier/noun frequency information can be used during online processing. We hope to address this in more detail in future work.
noun is typically unacceptable to native speakers. Thus massifiers are an open and flexible class, and are therefore plausibly less predictive than count classifiers. Moreover, according to Erbaugh (1986), only around 3 percent of nouns referring to physically co-present entities appear with a non-generic classifier in conversation. So it is possible that, while specific count classifiers are more restrictive than massifiers, comprehenders are less accustomed to having this information available during referential selection.

The fact that the massifiers prompt a stronger influence on comprehension in the visual world, despite being less restrictive, could be explained by another salient difference between measure words and massifiers on the one hand and count classifiers on the other: mass classifiers tend to align more with external visual properties than count classifiers do. That is, were a substance like milk to be arranged in visually distinct forms, it would likely be best denoted by different massifiers, as in ‘a glass of milk’, ‘a carton of milk’, or ‘a puddle of milk’. A count object like ‘snake’, however, can be shown in several different poses and still be paired with the same classifier in Chinese. In other words, massifiers may be more informative with regard to the visual world paradigm we used in these studies. Nevertheless, the ontological differences we found support the observation that massifiers refer to external containment or configuration of an object, which may be transient properties, while classifiers refer to physical or functional properties internal to the object’s structure, which are permanent (Tai 1994).

![Figure 14.12 Proportion of looks to the classifier targets over time(ms) in both specific and generic conditions, aligned to noun onset](image)

3 Unless, of course, the snake is crushed into a substance or is a member of a collective group.
14.7 Conclusions

In these studies, we find rapid effects of measure words, massifiers, and count classifiers on sentence processing in English and in Chinese. We have demonstrated that these words are being processed in a syntactically relevant way, and not merely exerting a superficial influence on comprehension. Our data also reveal slight differences between mass and count in Chinese, where this distinction is not syntactically realized. However, these differences may be driven by the visual nature of the task, and may reflect a general pattern in the language where massifiers encode more visually salient information than classifiers. The gradient facilitation we find from English measure phrases to Chinese massifiers, and finally to Chinese count classifiers is in accord with cognitive intuitions as well as linguistic observations that these words may contribute different types and degrees of information, although further research is necessary to fully explicate the mechanisms of mass and count reference across languages.
References


Anttila, Arto, and Adams Bodomo. 2006. A Dagaare-English word list. MS. Stanford University and University of Hong Kong.


References

References


References

References


References


References


References


References

Investigations 26: 159–173.
Jurafsky, Daniel. 1996. ‘Universal tendencies in the semantics of the diminutive’. Language
72: 533–578.
Kaden, Klaus. 1964. Der Ausdruck von Mehrzahlverhältnissen in der modernen chinesischen
Sprache. Berlin: Akademie.
Karimi, Simin. 1990. ‘Obliqueness, specificity, and discourse functions: Râ in Persian’. Lin-
Keenan, Edward. 1976. ‘Remarkable Subjects in Malagasy’. In Charles Li (ed.). Subject and
Oriental Society.
Kihm, Alain. 2005. ‘Noun classes, gender, and the lexicon-syntax-morphology interface:
A comparative study of Niger-Congo and Romance languages’. In Guglielmo Cinque and
University Press.
Kilarski, Marcin. 2007. ‘Algonquian and Indo-European gender in a historiographic perspec-
tive’. Historiographia Linguistica 34: 333–349.
ifornia, Irvine.
Köpcke, Klaus-Michael. 1982. Untersuchungen zum Genusstystem der deutschen Gegenwarts-
sprache. Tübingen: Niemeyer.
Kratzer, Angelika. 2008. ‘On the plurality of verbs’. In J. Dölling, T. Heyde-Zybatow, and
M. Schäfer (eds.). Event structures in linguistic form and interpretation, 269–301. Berlin:
Mouton de Gruyter.
Križka, Manfred. 1989. ‘Nominal reference, temporal constitution and quantification in event
semantics’. In J. van Benthem, R. Bartsch, and P. von Emde Boas (eds.). Semantics and
—— 1995. ‘Common nouns: A contrastive analysis of English and Chinese’. In G. Carlson and


References 295

Li, Peggy, and Lila Gleitman. 2002. ‘Turning the tables: Language and spatial reasoning’. 
Cognition 83: 265–294


MacDonald, Jonathan. 2008. The syntactic nature of inner aspect. Amsterdam: John Benjamins


—— 2001. ‘Words and things’. MS. MIT.


References

—— 2005b. ‘A part of wood is not a tree. On the absence of the count/mass distinction in Halkomelem’. Talk presented at the 40th International Conference on Salish and Neighbouring Languages, Musqueam Indian Band & University of British Columbia.
Yang, Defeng. 1996. ‘Liangci-qian shuci yi de yinxian wenti [The issue of the overtness of yi before a classifier]’. In Selected Papers of the Fifth Conference of the Association of Teaching Chinese as a Foreign Language. Beijing: Yuyan Xueyuan Press.
This page intentionally left blank
Index

accomplishment 129–30, 141–2
achievement 129–30, 142
acquisition 238–9, 241, 248–9, 257–9, 260
activity 129–30, 140–1, 145
agglomeration 48, 50
aggregate term 15–16
Agree 178
agreement 10, 11, 19, 25, 33, 78, 79 n., 123, 161, 175, 177–9, 182–3, 195, 239, 248–9
Akan 87
Aktionsart 129, 159, 192, see also aspect
Algonquian languages 163–4, 168 n., 169, 173, 174, 175, 177 n., 191, 193, 195, see also Blackfoot; Fox; Ojibwe
ambiguity 101–2
animacy 3, 4, 39, 41, 64, 81, 82, 83, 162–9, 170, 173–4, 177, 191–2, 193–5
Arabic 60, 109 n., 174, 195
Armenian 37, 239, see also Western Armenian
inner 147, 158, 162–3, 169–70
atomic parts 15, 23, 25, 247, 249, 254
atomicity 57 n., 80, 94, 119, 131, 132, 144–5, 204, 210–12, 214, 224, 234, 246
atomless term 15, 16

Bantu languages 6, see also Lingala; Swahili
Banyun 189
bare nominal 5, 6, 22, 32, 33, 42–3, 50, 55, 59, 60, 61, 64, 66, 100–10, 113–14, 121, 149, 150, 156, 170, 181–2, 185 n., 199, 205–7, 218–19, 239–40, 241, 244–8, 249, 252, 254, 255, 256–9
Bavarian 62–3, 67
Bayso 99–100

Biblical English 173, 189
Blackfoot 36, 64, 146–8, 151–7, 160–1, 162–70, 174, 191–2
borrowing 86–7, 92
boundedness 11 n., 140, 142, 147, 148, 154–5, 158–9, 161, 162–70, 191–2, 205–6
Breton 174, 181 n., 195–6, 216
Burmese 66

Cantonese 35 n., 37–8, 106, 199–219, 264
Cape Verdean Creole 110, 111
Catalan 103, 109–10, 111
change of state 192
Chechen 132, 143–4
Chinese languages 2, 6, 26 n., 30, 40 n., 55, 107, 159, 197, see also Cantonese; Hokkien; Mandarin
collective 227, 230, 234, 235, 236
cocurrence frequency 279–80
dividier 211–12, 213, 214
furniture nouns, associated with 210–11, 212, 214
generic 261–2
individual 220, 227, 228, 230, 233, 234, 235, 236, 237
individuating 211, 220, 226, 228, 229, 230, 234, 235, 236
kind 229, 234, 235, 236
language 5, 6, 23, 24, 25, 28, 39, 55, 199, 201, 227, 238–9, 240, 250, 252, 254, 255, 256, 257
linguistic relativity 264
classifier (cont.)
  mass, see massifier
  numeral 42, 52–3, 59 n., 64–6, 197, 220–37
  partitive 226–7, 230, 234, 235, 236
  phrase (ClP) 34–6, 38, 41–2
  plural 215–17
  shape 263–5
coercibility 3
  coercion 19, 20, 22, 25, 26, 30, 32, 57, 68–71,
  74, 116–17, 185, 218, 219, 241–4, 256, see
  also non-coerced mass noun
  collective 3, 43, 54–75, 76, 82, 90, 91, 92, 184,
  186, 191, 193, 194, 195–6, 223, 265, 270,
  see also classifier, collective
  compound 49–50, 86, 150, 157, 170, 175,
  179–80, 181 n., 182, 183, 202 n., 224
  constructional feature 17, 19, 22, 24, 25
  container-phrase 209
contextual inflectional morphology 190
  contrast 11, 23, 28–9, 30, 46, 81, 116, 122, 131,
  133, 141, 143, 157, 177 n., 191, 220–1, 222–4,
  225, 226, 228, 229, 231–2, 233, 235, 236,
  237, 239, 241, 244, 256
  contrastive specification 28–9, 30, 32, 39
conventionalized unit 117, 185
  count (nouns) (Chap3), 220–37
  countability 21, 25, 38, 58, 80, 156–8, 170, 220–37
  counting 6, 10, 26, 27, 38, 52, 59, 156, 173, 189,
  197, 199–219, 227, 229, 233, 234, 237, 247,
  249, 279
  construction 54, 58, 59, 64, 65–6, 67,
  72, 233
  Creole languages 110, see also Cape Verdean
  Creole; Haitian Creole; Singlish
cumulativity 1, 12, 140, 143
Dagaare 75–98, 195
  definite 39, 40–1, 48, 49, 50, 100, 108, 110, 111,
  112, 113–14, 115, 116, 117, 118, 119, 120–2,
  125, 127, 128, 205–6, 244, 245, 257, see also
  indefinite
  noun phrase (nP_{def}) 124, 126, 127
phrase (DefP) 40–2, 52
definiteness 5, 49, 50, 61, 112–13, 116, 118–20,
  121–2, 124, 128, 205–6, 216, 226
deixis 36
demonstrative 78, 101, 106, 123, 177, 178, 215,
  230, 257, 261, 271
Dene Syuline 23, 231, see also Slave
denominal verb 150, 170
dependent 28–9, 33, 35, 36, 37, 40–1, 47, 48,
  50, 52
derivational 86, 87, 137, 175, 176, 180–1, 182,
  183, 190, 196
determiner 5, 18–19, 24, 28, 29, 80, 100, 101,
  106–7, 110, 111, 148, 151, 152, 156, 170,
  177–8, 182, 256, 259, 262–3, 271, 278
  strong 44
dimension 223–7, 228, 229–30, 231–2,
  237, 265
  of individuation 29, 36
diminutive 89–90, 141, 144–5, 161
  Distributed Morphology (DM) 28, 124,
  126, 176
distributed predicate 43
distributive 44, 132, 133–4, 143, 144, 208, 239
  sub entailment 43–44
distributivity marker 43, 45
divisibility 1, 16, 57 n.
division 197
divisiveness 12, 16
dual 29 n., 82, 85, 91
  life 13–14, 18
  meaning 24
  noun 20
  use 14, 23, 24–5
Dutch 181 n., 190, 223

English 1–2, 4, 5, 6, 10, 11, 16, 21, 23, 26, 28,
  29–34, 35, 36, 37, 38, 39 n., 43, 45, 46, 47,
  52, 54, 55–6, 58, 60–1, 65, 66, 67–8, 71, 73,
  76, 79, 83, 91, 93–4, 96, 97, 104, 107, 113,
  116, 118, 122, 123, 126 n., 129 n., 144, 146,
  148–51, 154, 156–7, 158, 159–60, 162, 163,
  169, 170, 172, 173, 175, 176, 177, 179, 180,
- ensemble 19
- equative reading 49
- eventology 10 n.
- Ewe 90
- existential 39–40, 46, 49, 119, 257
- externalism 10
- eye-tracking 69 n., 264, 267, 275
- fact of language 16
- feature geometry 28, 128
- feature placing term 26
- Finnish 144
- Fox 168, 174, 193, 194
- French 6, 22, 44 n., 107, 173, 189, 191 n., 212 n., 246, 262, 263
- Frisian 92, see also West Frisian
- functional head 123, 124, 128, 173, 178, 183, 241
- shift 174, 191, 193–5, 198
- processing of 262, 263
- general number, see number, general
- German 56, 62, 65–6, 67, 70–1, 73, 146, 161–2, 163–4, 166–8, 192, 225, see also Bavarian; Middle High German
- Germanic languages 5, 70 n., see also Dutch; English; Frisian; German
- Greek 29, 117
- grinding 25, 218, see also Universal Grinder
- Gurenc 79
- Haitian Creole 110, 111, 112 n.
- head 27, 28, 29, 31, 32, 33, 34, 35, 40, 41, 46, 47, 50, 51, 52, 61, 101, 113, 114, 123–4, 126, 134, 135, 136, 144, 150, 176, 179, 180, 181, 233, see also functional head
- feature 27, 28
- Hebrew 29, 109, 173, 189
- Hokkien 106
- homogeneous in reference 12
- Hopi 2, 174–5, 227
- human 39, 52, 78, 79, 81, 86, 87, 174, 255, 258
- Hungarian 27, 62, 66, 126 n., 197, 216
- immaterial term 224, 228, 229
- inanimate, see animacy
- inclusiveness 118
- indefinite 10, 30, 32, 46, 47, 49, 50, 60, 61, 62, 67, 72, 100, 104, 110, 111, 113–14, 115, 116, 120, 122, 124, 125, 126, 127, 206, 216, 249, 267
- Indo-European languages 23, 55, 75, 146, 163–4, 166, 168, 169, 172, 173, 174, 175, 189
- Indonesian 104, 109, see also Riau
- Indonesian
- inflectional 29, 37, 44 n., 52, 67 n., 122, 123, 137, 157, 173, 174–83, 189, 190, 195
- inherent inflectional morphology 190
- inner aspect, see aspect, inner
- internalism 10
- Inuktitut 29
- Italian 107
- Javanese 104, 109
186–8, 197, 204 n., 205, 206, 208, 209 n., 215, 216, 217 n., 220, 222–3, 225, 226, 227, 230, 231, 234, 237, 238, 239, 246, 249, 252, 253–6, 261, 262, 271, 274, 279, see also classifier, numeral

object-mass noun 55, 245–7, 254, 255, 256
obviative 179 n., 194–5
Ojibwe 71 n., 151 n., 172–98
Old Persian 67–8
ontological distinction 3, 17, 26, 59, 147–8, 150, 155, 240–1, 259, 262, 263, 279

parameter 21–2, 107, 108, 173, 175, 176, 188, 196

pejorative 141, 145
Persian 5, 45–52, 55–6, 61, 62, 63, 66, 67, 69, 73, 111, 112–28, 173, 189, 190, see also Old Persian
person 36, 40, 41, 101, 178, 179, 194, 221
person/gender hierarchy 179 n.
Pirahâ 100
plurational 129, 131, 132 n.

optionality of 112, 113, 116, 126–7
plural of abundance 189
plurality tantum 55 n., 123, 181, 182
plurality 2, 6, 7, 27, 36–7, 38, 39, 40 n., 45, 48, 52, 82, 85, 86, 92, 94, 95, 126 n., 131, 132, 143, 144, 215, 217, 223, 244, 251
pluralization 1–2, 5, 6, 18 n., 24, 54, 55, 57 n., 58, 63, 64, 66, 67, 71, 123, 141, 144, 149, 150, 153, 156, 160, 167, 170, 172, 173, 177, 179, 180, 181, 182, 193, 196, 239, 240

of mass nouns 117, 119, 172, 173, 175, 176, 183–91, 192, 195, 198
Polish 60, 81
portmanteau morpheme 121, 125
precisifier 128
privative 28
projection 22, 27, 28, 29, 30, 32, 33, 34, 35, 40 n., 42, 44, 47, 50, 52, 113, 122, 123, 128, 175, 182, 197, 208, 233
quanta 29, 30
quantifier 10, 13, 46, 80, 101 n., 108, 116, 135, 148, 151, 156, 159, 161, 170, 183, 184, 186–8, 221, 226, 239, 240, 249, 254, 257, see also universal quantifier
quantity 7, 32, 38, 46, 54, 58, 59, 117, 121, 202 n., 203, 211, 212, 214, 219, 256, 265, see also Maxim of Quantity
quantization 7, 140, 143
quantize 29, 140
reclassification 149, 153
redundancy avoidance 126, 127
register 47
Riau Indonesian 106
Romance languages 2, 5, 109–10, see also Catalan; French; Italian; Spanish
root 95, 124, 131, 132, 135, 136–9, 141, 143, 144, 148, 150, 151, 157, 158, 159, 164, 168, 170, 173, 175, 176, 179, 180, 181, 193, 231
Russian 144, 163 n., 166 n.

Salish languages, see Halkomelem;
Lillooet; Squamish
scope 44, 104–6, 109, 110, 111
of contrast 28
Seinsart 147, 159, see also aspect
selectability 148, 151, 168
selection 147, 221
semantic anomaly 12, 13, 19
semantic compositionality 19, 25
semantic feature 10, 13, 17, 21
semantic interpretability constraint 13
semi-lattice 15, 19, 20, 37 n., 94, 95, 96, 97, 247
shape 66, 220, 223, 224, 232, 237, 250, 251, 252, 263, 264–5
Singaporean Malay 106
Singlish 106
singularization 191, 193, 209
singulative 88–9, 91, 92, 173, 184 n., 191, 193–4, 195–7, 198
size 20, 81, 154, 211 n., 220, 223, 224–5, 232, 237
Slave 174, see also Dene Sulíne
Slavic languages 107, see also Polish; Russian
sortal 26, 200 n., 202 n., 240, 241, 242, 253, 259
Spanish 73, 109, 263
specifier 32, 34, 35, 38, 46, 47
Squamish 129 n., 131, 132, 135
stage 129–30, 140, 141, 142
state 29, 129–30, 192
supervaluation 23, 25
Swahili 60, 61, 91
syntactic ill-formedness 13, 17, 19, 25
syntagmatic 220–5, 231
Tagalog 197
Taos 175
taxonomic reading 115, 117
telicity 129–31, 132 n., 139–43, 150, 169, 192
Thai 6
time 36, 129, 191
totality reading 116

transnumeral 59–60, 61–4, 65, 66–7, 68–70, 71, 72, 73
Turkish 63, 64, 66, 104, 109, 197, 253
type-shift 20, 22, 23, 25
unbounded event, see boundedness
underspecified 22, 25, 106, 110 n., 247, 254
unit marker 208, 213, 219
unit word 221, 227, 229–31, 233, 235, 236, 237
Universal Grinder 2, 4, 14, 20, 29, 30, 185, 199–200, 217–18, 219
Universal Packager 14, 20, 185, 227
universal quantifier 105, 257, 261
reading 116
Universal Sorter 29, 30, 185, 227
unspecified for mass/count 18, 21, 23, 24, 30
vagueness 23, 25, 28, 38, 58 n., 128, 190

Welsh 91, 92, 195
West Frisian 92–3
Western Armenian 252, 253–9
Yudja 23, 227, 231
Zuni 175
OXFORD STUDIES IN THEORETICAL LINGUISTICS

Published
1 The Syntax of Silence
Sluicing, Islands, and the Theory of Ellipsis
by Jason Merchant
2 Questions and Answers in Embedded Contexts
by Utpal Lahiri
3 Phonetics, Phonology, and Cognition
edited by Jacques Durand and Bernard Laks
4 At the Syntax-Pragmatics Interface
Concept Formation and Verbal Underspecification in Dynamic Syntax
by Lutz Marten
5 The Unaccusativity Puzzle
Explorations of the Syntax-Lexicon Interface
edited by Artemis Alexiadou, Elena Anagnostopoulou, and Martin Everaert
6 Beyond Morphology
Interface Conditions on Word Formation
by Peter Ackema and Ad Neeleman
7 The Logic of Conventional Implicatures
by Christopher Potts
8 Paradigms of Phonological Theory
edited by Laura Downing, T. Alan Hall, and Renate Raffelsiefen
9 The Verbal Complex in Romance
by Paola Monachesi
10 The Syntax of Aspect
Deriving Thematic and Aspectual Interpretation
Edited by Nomia Eretschik-Shir and Tova Rapoport
11 Aspects of the Theory of Clitics
by Stephen Anderson
12 Canonical Forms in Prosodic Morphology
by Laura J. Downing
13 Aspect and Reference Time
by Olga Borik
14 Direct Compositionality
edited by Chris Barker and Pauline Jacobson
15 A Natural History of Inflection
by Alan C. L. Yu
16 Phi-Theory
Phi-Features Across Interfaces and Modules
edited by Daniel Harbour, David Adger, and Susana Béjar
17 French Dislocation: Interpretation, Syntax, Acquisition
by Cécile De Cat
18 Inflactional Identity
edited by Asaf Bachrach and Andrew Nevis
19 Lexical Plurals
by Paolo Acquaviva
20 Adjectives and Adverbs
Syntax, Semantics, and Discourse
Edited by Louise McNally and Christopher Kennedy
21 InterPhases
Phase-Theoretic Investigations of Linguistic Interfaces
edited by Kleanthes Grohmann
22 Negation in Gapping
by Sophie Repp
23 A Derivational Syntax for Information Structure
by Luis López
24 Quantification, Definiteness, and Nominalization
edited by Anastasia Giannakidou and Monika Rathert
25 The Syntax of Sentential Stress
by Arsalan Kahemuyipour
26 Tense, Aspect, and Indexicality
by James Higginbotham
27 Lexical Semantics, Syntax and Event Structure
edited by Malka Rappaport Hovav, Edit Doron and Ivy Sichel
28 About the Speaker
Towards a Syntax of Indexicality
by Alessandra Giorgi
29 The Sound Patterns of Syntax
edited by Nomia Eretschik-Shir and Lisa Rochman
30 The Complementizer Phase
edited by Phoevos Panagiotidis
31 Interfaces in Linguistics
New Research Perspectives
edited by Raffaella Folli and Christiane Ulbrich
32 Negative Indefinites
by Doris Penka
33 Events, Phrases, and Questions
by Robert Truswell
34 Dissolving Binding Theory
by Johan Rooryck and Guido Vanden Wyngaerd
35 The Logic of Pronominal Resumption
by Ash Asudeh
36 Modals and Conditionals
by Angelika Kratzer
37 The Theta System
Argument Structure at the Interface
edited by Martin Everaert, Marijana Marej, and Tal Siloni
38 Sluicing
Cross-Linguistic Perspectives
edited by Jason Merchant and Andrew Simpson
39 Telicity, Change, and State
A Cross-Categorial View of Event Structure
edited by Violeta Demonte and Louise McNally
40 Ways of Structure Building
edited by Myriam Uribe-Etxebarria and Vidal Valmala
41 The Morphology and Phonology of Exponence
edited by Jochen Trommer
42 Count and Mass Across Languages
edited by Diane Massam
43 Genericity
*edited by* Alda Mari, Claire Beyssade, and Fabio Del Prete

44 Strategies of Quantification
*edited by* Kook-Hee Gil, Steve Harlow, and George Tsoulas

**Published in association with the series**
The Oxford Handbook of Linguistic Interfaces
*edited by* Gillian Ramchand and Charles Reiss

**In preparation**
The Syntax of Roots and the Roots of Syntax
*Edited by* Artemis Alexiadou, Hagit Borer, and Florian Schäfer

External Arguments in Transitivity Alternations
*by* Artemis Alexiadou, Elena Anagnostopoulou, and Florian Schäfer

Semantic Continuations
Scope, Binding, and Other Semantic Side Effects
*by* Chris Barker and Chung-Chieh Shan

Phi Syntax: A Theory of Agreement
*by* Susana Béjar

Stratal Optimality Theory
*by* Ricardo Bermúdez Otero

Diagnosing Syntax
*edited by* Lisa Lai-Shen Cheng and Norbert Corver

Phonology in Phonetics
*by* Abigail Cohn

Generality and Exception
*by* Ivan Garcia-Alvarez

Focus and Ellipsis
*by* Kirsten Gengel

The Indefiniteness and Focusing of Wh-words
*by* Andreas Haida

Edges and Cyclic Syntax: Evidence from Scrambling in Korean and Japanese
*by* Heejeong Ko

The Semantics of Evaluativity
*by* Jessica Rett

Computing Optimality
*by* Jason Riggle

Nonverbal Predications
*by* Isabelle Roy

Null Subject Languages
*by* Evi Sifaki and Ioanna Sitaridou

Gradience in Split Intransitivity
*by* Antonella Sorace