

Aikhenvald
Classifiers

Oxford
LINGUISTICS

Classifiers

*A Typology of Noun Categorization
Devices*

Alexandra Y. Aikhenvald

Oxford Studies in Typology and Linguistic Theory

LI7.
categorically
category /'k
its source la

OXFORD

CLASSIFIERS

OXFORD STUDIES IN TYPOLOGY AND LINGUISTIC THEORY

SERIES EDITORS: Ronnie Cann, *University of Edinburgh*, William Croft, *University of Manchester*, Mark Durie *University of Melbourne*, Anna Siewierska, *University of Lancaster*

This series offers a forum for original and accessible books on language typology and linguistic universals. Works published will be theoretically innovative and informed and will seek to link theory and empirical research in ways that are mutually productive. Each volume will also provide the reader with a wide range of cross-linguistic data. The series is open to typological work in semantics, syntax, phonology, and phonetics or at the interfaces between these fields.

Published:

Indefinite Pronouns

Martin Haspelmath

Intransitive Predication

Leon Stassen

Classifiers: A Typology of Noun Categorization Devices

Alexandra A. Aikhenvald

Anaphora

Yan Huang

In preparation:

The Noun Phrase

Jan Rijkhoff

Double Object Constructions

Maria Polinsky

CLASSIFIERS

A Typology of Noun Categorization Devices

ALEXANDRA Y. AIKHENVALD

OXFORD
UNIVERSITY PRESS

OXFORD

UNIVERSITY PRESS

Great Clarendon Street, Oxford OX2 6DP

Oxford University Press is a department of the University of Oxford.
It furthers the University's objective of excellence in research, scholarship,
and education by publishing worldwide in

Oxford New York

Athens Auckland Bangkok Bogotá Buenos Aires Calcutta
Cape Town Chennai Dar es Salaam Delhi Florence Hong Kong Istanbul
Karachi Kuala Lumpur Madrid Melbourne Mexico City Mumbai
Nairobi Paris São Paulo Singapore Taipei Tokyo Toronto Warsaw
and associated companies in Berlin Ibadan

Oxford is a registered trade mark of Oxford University Press
in the UK and certain other countries

Published in the United States
by Oxford University Press Inc., New York

© A. Y. Aikhenvald 2000

The moral rights of the author have been asserted
Database right Oxford University Press (maker)

First published 2000

All rights reserved. No part of this publication may be reproduced,
stored in a retrieval system, or transmitted, in any form or by any means,
without the prior permission in writing of Oxford University Press,
or as expressly permitted by law, or under terms agreed with the appropriate
reprographics rights organizations. Enquiries concerning reproduction
outside the scope of the above should be sent to the Rights Department,
Oxford University Press, at the address above

You must not circulate this book in any other binding or cover
and you must impose the same conditions on any acquirer

British Library Cataloguing in Publication Data
Data available

Library of Congress Cataloging in Publication Data
Data applied for

ISBN 0-19-823886-X

1 3 5 7 9 10 8 6 4 2

Typeset in Times
by J&L Composition Ltd, Filey, North Yorkshire
Printed in Great Britain
on acid-free paper by
Biddles Ltd, Guildford & King's Lynn

For Bob, okojibotee

This page intentionally left blank

Preface

This book aims at providing a cross-linguistic analysis of noun classification systems across the languages of the world, also dealing with a variety of other problems such as the morphological status of the markers of these categories, agreement phenomena, and the syntactic and semantic classification of adjectives and numbers. It is generally accepted that linguistic categorization of nouns is a reflection of human mind and culture. The present study thus has far-reaching implications for cross-cultural as well as cross-linguistic studies of human cognition, and will provide new insights concerning the mechanisms by which human language functions.

Languages with extensive systems of noun classification devices, especially those which combine classifiers and genders, present a true challenge for the typologist. My first encounter with these unusual systems was through fieldwork on Tariana and Baniwa, two closely related North Arawak languages spoken in Northwest Amazonia. The more I worked on the topic, the more exotic and unusual systems I encountered, especially among little-known South American languages, and languages of the South Pacific. This book came into being as an attempt to integrate these systems into a cross-linguistically based typological framework.

This study is an up-to-date introduction to the field, and will be of value not only to a wide variety of linguists and linguistic students but also to anthropologists, cognitive psychologists, and philosophers who are interested in language and the mind. It can be used both as a sourcebook for further typological studies, and as a textbook. The discussion in the book is in terms of basic linguistic theory, the framework of linguistic analysis in terms of which most grammars are cast, and in terms of which significant typological generalizations are postulated. (I have avoided using any of the more specific formalisms, which come and go with such frequency.)

Some terminological clarifications are in order. First, my conception of a lexical entry for 'noun' roughly corresponds to the notion of 'lexeme' as outlined by Lyons (1977 vol. 1: 19). Second, throughout the book 'linguistic categorization of a noun' is used to mean 'linguistic categorization of the referent of a noun', just as in many linguistic usages 'human noun' is a short way of saying 'noun with a human referent'. Third, the term 'noun categorization' is used here in a sense close to the 'noun classification' (cf. Craig 1986a; Derbyshire and Payne 1990) or 'nominal classification' (cf. Harvey and Reid 1997) employed by other authors. The term 'classifier system' refers to a grammatical system of noun categorization device(s) in a particular language.

In order to limit the book to a reasonable size, I have only been able to refer to a portion of the available literature. There are many other sources that I have consulted, which only provide additional exemplification for points that are already well covered. When a language is introduced for the first time, its genetic affiliation and the source of information on it are given in parentheses; further on, this information is only repeated where relevant. Examples, tables and diagrams are numbered separately within each chapter.

The orthography used in the examples and language names follows that of the sources (unless indicated otherwise).

A study like this could only be definitive when good and thorough descriptions have been provided for most of the world's languages; we are at present a long way from this situation. Nevertheless, I hope that this study will provide a framework within which fieldworkers and typologists will be able to work, and which can be amended and adjusted as new data and new insights emerge.

It is my hope that this book will encourage people to study noun classification devices, especially in little-known or undescribed languages, going out into the field and documenting languages threatened by extinction (before it is too late to do so).

Acknowledgements

My gratitude goes to all those native speakers who taught me their languages and their unusual classifier and gender systems: Cândido, José, Jovino, Graciliano, and Olívia Brito (Tariana); Humberto Baltazar and Pedro Ângelo Tomas (Warekena); the late Candelário da Silva (Bare); Afonso, Albino and João Fontes, Celestino da Silva, Cecília and Laureano da Silva, and the late Marcília Rodrigues (Baniwa); the late Tiago Cardoso (Desano, Piratapuya); Alfredo Fontes (Tucano); Marilda and Carlito Paumari (Paumari); Raimunda Palikur (Palikur); Simone Nientao (Tamachek) and—last but not least—Pauline and James Laki (Manambu).

I am also indebted to students in the Federal University of Santa Catarina, Brazil, and in the Australian National University. I learned a lot from working with Rute Amorim, Lílias Chun, Christiane Cunha de Oliveira, Tim Curnow, Michael Dunn, Catriona Hyslop, Dorothy Jauncey, Yunseok Lee, Eva Lindström, Peita Littleton, Rina Marnita, Silvana Martins, Kazuko Obata, Kristina Sands, Eva Tatrai, Angela Terrill, Simoni Valadares, and Jacki Wicks. My warmest thanks go to Silvana and Valteir Martins and Lenita and Elias Coelho de Assis, without whose friendship and assistance a great deal of my fieldwork would have been impossible. Special gratitude goes to Diana Green, who revealed to me the beauty of genders and classifiers in Palikur.

I am most grateful to those people who helped me by sending copies of their papers, answering my questions and commenting on various parts of this manuscript: José Alvarez, Mengistu Amberber, Felix Ameka, Peter van Baarle, Janet Barnes, Cândida Barros, Edith Bavin, Walter Bisang, Kim Blewett, Paula Boley, John Boyle, Friederike Braun, Lea Brown, Bill Callister, Lyle Campbell, Eugene Casad, Meiyun Chang-Smith, Adam Chapman, Shirley Chapman, Hilary Chappell, Helen Charters, Bernard Comrie, Bob Conrad, Grev Corbett, Tim Curnow, Des Derbyshire, Connie Dickinson, Tony Diller, Gerritt Dimmendaal, Mark Donohue, Nancy Dorian, Mark Durie, Tom Dutton, Nora England, Nick Evans, Cindi and Jim Farr, Bill Foley, Lys Ford, David Foris, Paul Frank, David Gil, Cliff Goddard, Elsa Gomez-Imbert, Ian Green, Rebecca Green, Colette Grinevald (Craig), Geoff Haig, Mark Harvey, Rie Hasada, Bernd Heine, Debbie Hill, Chu-Ren Huang, Rodney Huddleston, Suanu Ikoro, Liisa Järvinen, Jae Jung Song, Aleksandr J. Kibrik, Harold Koch, Antonina I. Koval', Randy LaPolla, Jason Lee, Jennie Lee, Jeff Leer, W. P. Lehmann, Adrienne Lehrer, Frank Lichtenberk, Eva Lindström, Elizabeth Löbel,

Ivan Lowe, Harriet Manelis-Klein, Jack Martin, Marianne Mithun, Catherine McGuckin, Ulrike Mosel, Otto Nekitel, Bee Chin Ng, Johanna Nichols, Masayuki Onishi, Patricia Pacioni, Helma Pasch, Peter Paul, Andrew Pawley, David Payne, Vladimir Plungian, Bill Poser, Kostantin Pozdnjakov, Bruce Rigsby, Phil Quick, Susan Quigley, Katya Rakhilina, Henri Ramirez, Nick Reid, Keren Rice, Aryon Rodrigues, Malcolm Ross, Carl Rubino, Alan Rumsey, Filomena Sândalo, Risto Sarsa, Hansjakob Seiler, Lucy Seki, Gunter Senft, Beatriz and Rodolfo Senn, Gi-Hyun Shin, Tim Shopen, Edgar Suter, Chad Thompson, Irina Toporova, Joe Tsonope, Ione Vasconcelos, Alejandra Vidal, Tiit-Rein Viitso, Viktor A. Vinogradov, Julie Waddy, Bruce Waters, Laurel Watkins, Anna Wierzbicka, Mary Ruth Wise, Stephen Wurm, and Roberto Zavala.

I am most grateful to those who read through the whole draft of this book, or parts of it, and provided comments, corrections and ideas—Peter Denny, Nancy Dorian, David Foris, Cliff Goddard, Nikolaus Himmelmann, Frank Lichtenberk, Edith Moravcsik, Patricia Pacioni, Helma Pasch, Doris Payne, Nick Reid, Malcolm Ross, Fritz Serzisko, and Roberto Zavala. Invaluable comments on almost every page came from R. M. W. Dixon, Keith Allan, Walter Bisang, Lyle Campbell, Gerritt Dimmendaal, Mark Durie, Ulrike Mosel, and Gunter Senft.

My deepest gratitude goes to the SIL Library in Ukarumpa, and to Paul Frank, the director of the Summer Institute of Linguistics (Colombia), who provided me with invaluable materials on languages of Papua New Guinea and Colombia respectively.

I am also grateful to the members of Eesti Noorte Grupp of Canberra—Reet Bergman, Krista Gardiner, and Reet Vallak—who helped me realize that communicating in a language without genders or classifiers can be great fun.

Suzanne Kite carefully read through several drafts of this book and corrected it with her usual skill, dedication, and good humour. Thanks are equally due to her.

Jennifer Elliott provided a wonderful working atmosphere at the Research Centre for Linguistic Typology. This book would have been scarcely possible without her.

Plea

This book is far from being the last word on noun categorization devices. I welcome reactions, counterexamples, new ideas and data, to further develop, refine, and improve the generalizations put forward here. Please send them to me at Research Centre for Linguistic Typology, La Trobe University, Bundoora Vic., 3083 Australia.

This page intentionally left blank

Contents

List of Maps	xx
List of Tables	xxi
List of Diagrams	xxiv
List of Abbreviations	xxv
1. Preliminaries	1
1.1 General remarks	1
1.2 Classifiers: an illustration	1
1.3 Theoretical framework, data, and sources	4
1.4 Approaches to the typology of classifiers	5
1.5 Parameters for the typology of classifiers	13
1.6 The structure of this book	16
2. Noun Class and Gender Systems	19
2.1 General remarks	19
2.2 Properties of noun class systems	20
2.3 Principles of noun class assignment	22
2.3.1 Semantic assignment	22
2.3.2 Morphological assignment	25
2.3.3 Phonological assignment	25
2.3.4 Mixed principles of assignment	25
2.4 Noun classes and agreement	28
2.4.1 A working definition of agreement and agreement properties	28
2.4.2 Principles of noun class agreement	31
2.4.3 Variability in noun class assignment and variable agreement	41
2.4.4 Determining the number of noun classes in a language	45
2.5 Markedness and resolution in noun classes	50
2.5.1 Markedness	50
2.5.2 Noun class resolution	52
2.5.3 Markedness relationships in noun classes	54
2.6 Realization of noun classes	56
2.6.1 Overt and covert noun class marking	57
2.6.2 Morphological realization of noun classes	58
2.6.3 Double marking of noun classes	63

2.7	Languages with more than one kind of noun class	67
2.7.1	Nominal and pronominal noun class	68
2.7.2	Different kinds of noun class in the same environment	70
2.7.3	Languages with more than one kind of noun class: a summary	76
2.8	Distribution of noun classes in the languages of the world	77
3.	Noun Classifiers	81
3.1	Properties of noun classifiers	81
3.2	Noun classifiers: discussion and exemplification	82
3.2.1	The choice of noun classifiers and the cooccurrence of several classifiers within one noun phrase	82
3.2.2	Semantic functions of noun classifiers	84
3.2.3	Size of inventory and degree of grammaticalization of noun classifiers	84
3.2.4	Syntactic functions of noun classifiers	87
3.3	Noun classifiers and numeral classifiers	90
3.4	Realization and grammaticalization of noun classifiers	91
3.5	Overt noun class marking and noun classifiers	92
3.6	Distribution of noun classifiers in the languages of the world	97
4.	Numeral Classifiers	98
4.1	Properties of numeral classifiers	98
4.2	Numeral classifier constructions and morphological realization of numeral classifiers	101
4.2.1	Numeral classifiers as independent lexemes	101
4.2.2	Numeral classifiers attached to numerals	105
4.2.3	Numeral classifiers attached to the head noun	110
4.3	Languages with more than one morphological type of numeral classifier	112
4.3.1	Different types of numeral classifier in complementary distribution	112
4.3.2	Different types of numeral classifier which occur together	113
4.4	Problems with numeral classifiers	114
4.4.1	Mensural and sortal classifiers: distinguishing classifiers from quantifying expressions	114
4.4.2	Incipient numeral classifiers	120
4.5	Distribution of numeral classifiers in the languages of the world	121

5. Classifiers in Possessive Constructions	125
5.1 Categorization in possessive constructions	125
5.2 Possessed classifiers	126
5.3 Relational classifiers	133
5.3.1 Relational classifiers and their properties	133
5.3.2 Types of possession and relational classifiers	137
5.4 Possessor classifiers	139
5.5 Interaction of possessed and relational classifiers	140
5.5.1 Integrating relational and possessive classifiers	140
5.5.2 Languages with two types of classifier in possessive constructions	142
5.6 Contrasting classifiers in possessive constructions	144
5.7 Distribution of classifiers in possessive constructions in the languages of the world	147
6. Verbal Classifiers	149
6.1 Properties of verbal classifiers	149
6.2 Realization of verbal classifiers	149
6.2.1 Classificatory noun incorporation	150
6.2.2 Verbal classifiers as affixes	152
6.2.3 Suppletive 'classificatory verbs'	153
6.2.4 The interaction of the three types of verbal classifier	160
6.3 Verbal classifiers and syntactic function of the argument	162
6.4 Combinations of different types of verbal classifier	163
6.4.1 Different types of verbal classifier in complementary distribution	163
6.4.2 Distinct systems of verbal classifiers	167
6.5 Distribution of verbal classifiers	169
7. Locative and Deictic Classifiers	172
7.1 The structure of this chapter	172
7.2 Properties of locative classifiers	172
7.3 Properties of deictic classifiers	176
7.3.1 Examples of deictic classifiers	177
7.3.2 Conclusions and discussion	181
8. Different Classifier Types in One Language	184
8.1 General observations	184
8.2 Coexisting classifier sets in different environments	185
8.3 Different classifier sets in the same environment	198
8.4 Conclusions	201

9. Multiple Classifier Languages	204
9.1 Noun categorization in multiple classifier languages	204
9.2 Multiple classifier languages and noun class agreement on multiple targets	228
9.3 Fuzzy types: overlapping classifiers in multiple environments	230
9.3.1 Multiple classifiers in Baniwa	230
9.3.2 Multiple classifiers in Tariana	235
9.3.3 Fuzzy types and borderline cases	240
10. Classifiers and Other Grammatical Categories	242
10.1 Classifiers and number	243
10.1.1 Noun classes and number	243
10.1.2 Number and other classifier types	249
10.2 Classifiers and person	252
10.3 Classifiers and grammatical function	255
10.3.1 Noun classes and grammatical function	255
10.3.2 Verbal classifiers and grammatical function	257
10.4 Classifiers and types of possession	257
10.4.1 Noun classes and types of possession	258
10.4.2 Classifiers in possessive constructions and types of possession	259
10.5 Classifiers and politeness	260
10.6 Classifiers and declensional classes	262
10.7 Classifiers and verbal categories	263
10.8 Classifiers and deictic categories	266
10.9 Classifiers, derivation, and lexicon	266
10.10 Conclusions	268
11. Semantics of Noun Categorization Devices	271
11.1 Semantic parameters in noun categorization	271
11.1.1 Basic parameters of categorization	271
11.1.2 Additional semantic characteristics	274
11.1.3 Semantic relationship between a classifier and the referent	275
11.2 Semantics of classifier types	275
11.2.1 Semantics of noun classes	275
11.2.2 Semantics of noun classifiers	283
11.2.3 Semantics of numeral classifiers	286
11.2.4 Semantics of classifiers in possessive constructions	293
11.2.5 Semantics of verbal classifiers	295

11.2.6	Semantics of locative and of deictic classifiers	300
11.2.7	Semantic parameters in languages with several different types of classifier	301
11.2.8	Semantic parameters in multiple classifier systems	303
11.2.9	Conclusions	305
12.	Semantic Organization and Functions of Noun Categorization	307
12.1	Semantic organization and functions of classifier systems	307
12.1.1	Semantic complexity in classifier systems	308
12.1.2	Semantic roles of classifiers	317
12.1.3	Discourse-pragmatic functions of classifiers	320
12.1.4	Applicability of classifiers and default classes	334
12.2	Human cognition and classifiers	337
12.2.1	Perceptual correlates of noun categorization	337
12.2.2	Cognitive mechanisms and noun categorization	339
12.3	Social and cultural issues in noun categorization	340
12.3.1	Social structure in noun categorization	342
12.3.2	Environment and culture in noun categorization	343
12.3.3	Culture-specific metaphorical extensions	346
12.3.4	Socio-cultural motivations for change in noun categorization	347
12.4	Conclusions	350
13.	Origin and Development of Noun Categorization Devices	352
13.1	Lexical sources for classifiers	353
13.1.1	From a noun to a classifier	353
13.1.2	Repeater phenomena and the origin of classifier constructions	361
13.1.3	From a verb to a classifier	362
13.1.4	Classifiers from deverbal nominalizations	365
13.1.5	Classifiers of mixed origin	366
13.2	From a closed class to a noun categorization system	367
13.3	Languages with several classifier types, and the relative age of noun categorization devices	370
13.4	Internal evolution of noun categorization	372
13.5	Grammaticalization and reanalysis in noun categorization systems	374
13.5.1	Grammaticalization in the development of noun categorization	374
13.5.2	Reanalysis in noun categorization	377

13.6	Reduction and loss of noun categorization devices	379
13.7	Language-external motivations for the development and decay of noun categorization	382
13.7.1	Language contact and noun categorization	383
13.7.2	Creolization and noun categorization	388
13.7.3	Language obsolescence and noun categorization	389
13.7.4	Language-external motivations and their impact on noun categorization	391
13.8	Development and loss of agreement	391
13.8.1	The genesis and development of agreement	391
13.8.2	Decline and loss of agreement	398
13.9	Semantic changes in noun categorization devices	400
13.9.1	From lexical item to classifier: principles of semantic change	401
13.9.2	Further changes in noun categorization devices	407
13.10	Sources of noun categorization devices: a summary	411
14.	Noun Categorization Devices in Language Acquisition and Dissolution	413
14.1	Acquisition and development of noun classes	413
14.2	Acquisition of numeral classifiers	417
14.3	Dissolution of noun classes and of numeral classifiers	422
14.4	Conclusions	423
15.	Conclusions	425
15.1	Properties of classifier types	425
15.2	Cooccurrence of classifier types and multiple classifier languages; prototypes and continua	432
15.3	Prospects for future studies	434
Appendix 1.	Noun Categorization by Means Other than Classifiers	436
Appendix 2.	From Nouns to Classifiers: Further Examples of Semantic Change	442
(A)	Body parts as sources for classifiers: semantic extensions	442
(B)	Sources for shape-based numeral classifiers	446
Appendix 3.	Fieldworker's Guide to Classifier Languages	447
References		452
List of Languages		489
List of Language Families, Linguistic Areas, and Proto-languages		504

Index of Languages, Linguistic Areas, and Language Families	509
Index of Authors	519
Subject Index	525

List of Maps

1. Distribution of noun classes and genders in the languages of the world	78
2. Distribution of noun classifiers in the languages of the world	96
3. Distribution of numeral classifiers in the languages of the world	122
4. Distribution of classifiers in possessive constructions in the languages of the world	148
5. Distribution of verbal classifiers in the languages of the world	170

List of Tables

1.1	Shopping list in Japanese	2
1.2	Differences between noun classes and classifiers	6
1.3	Classifiers, their morphological realization, and semantics	7
1.4	Classifiers and their functions	8
2.1	Semantics of noun classes in Proto-Bantu	24
2.2	Semantic features for the gender assignment of inanimate nouns in Cantabrian Spanish	27
2.3	Inflection and derivation	30
2.4	Gender marking in Rumanian	46
2.5	Gender marking in Telugu	46
2.6	Gender agreement in Khinalug	47
2.7	A fragment of the Russian nominal paradigm	48
2.8	Noun classes in Ingush	49
2.9	Noun classes in Ndali	64
2.10	'Pronominal' and 'nominal' noun class systems	68
2.11	<i>Ka</i> -class assignment in Paumari	73
2.12	Two types of noun class (genders) in Mba	75
4.1	Numeral classifiers used with humans in Assamese	102
4.2	Numeral classifiers in Telugu	108
4.3	Classifiers fused with numerals in Kusaiean	108
4.4	Numeral classifiers in Nivkh	109
4.5	Numeral classifiers in Warekena	109
4.6	Numeral classifiers in Squamish	110
5.1	Possessed classifiers in Panare	128
5.2	Systems of two relational classifiers	134
5.3	A system of relational classifiers in Boumaa Fijian	134
5.4	Sample of classifiers in Puluwat	141
5.5	Differences between possessed and relational classifiers	145
5.6	Relational, possessed, and possessor classifier: a comparison	146
6.1	Mescalero Apache classificatory verb categories	155
6.2	Classificatory verbs in Ojibway	155
6.3	Classificatory verbs in Ika	156
6.4	Examples of the use of 'give' in Mescalero Apache	157
6.5	Classificatory verbs in Koyukon: an example	157
6.6	Classificatory verbs in Nevome	158
6.7	Classificatory verbs in Enga	159
6.8	Classificatory verbs in Ku Waru	159
6.9	Verbal classifiers in Waris	166

6.10	Classificatory verbs in Waris	167
6.11	Affixed verbal classifiers in Koyukon	167
7.1	Locative classificatory suffixes in three Carib languages	175
7.2	Article classifiers in Ponca	178
7.3	A sample of demonstratives in Proto-Guaicuruan	181
7.4	Deictic classifiers in Eskimo	182
8.1	Noun classes and noun classifiers in Ngan'gityemerri	186
8.2	Classifiers in Mokilese	187
8.3	Numeral classifiers in Akatek	188
8.4	Noun classifiers in Akatek	188
8.5	Numeral, verbal, and locative classifiers in Palikur	193
8.6	Demonstratives in Palikur (singular)	194
8.7	Gender marking on verbs in Palikur	194
8.8	Properties of classifiers and genders in Palikur	195
8.9	Semantic and functional properties and origin of classifiers in Palikur	197
8.10	Different classifier sets in different environments in one language	202
8.11	Different classifier sets in the same environment in one language	202
9.1	Same set of classifiers in several environments (A–D)	207
9.2	Locative and verbal classifiers in Eyak	209
9.3	Same morphemes in several classifier environments	225
9.4	Environments in which genders and classifiers are used in Baniwa and Tariana	230
9.5	Classifiers in Baniwa	232–4
9.6	Classifiers in Tariana	236–8
9.7	Agreement forms of <i>kwa-</i> in Tariana	240
10.1	Personal pronouns in Tamachek	245
10.2	Personal pronouns in Lithuanian	245
10.3	Personal pronouns in Slovene	246
10.4	Personal pronouns in Resígaro and in Bora	247
10.5	Gender in Malto	248
10.6	Classificatory verbs in Tewa and their semantics	251
10.7	Animacy marking on verbs in Jarawara (Arawá)	251
10.8	Personal pronouns in Minangkabau (singular)	253
10.9	Personal pronouns in Spanish	254
10.10	Personal pronouns in Tariana	254
10.11	Paradigm of Latin <i>is</i> 'this'	256
10.12	Possessed classifiers and speech styles in Ponapean	261
10.13	Examples of possessed classifiers in common and humiliative speech in Ponapean	261
10.14	Paradigm of <i>dela-t</i> 'do' in past tense in Russian	264
10.15	Paradigm of <i>katav</i> 'write' in Modern Hebrew	264
10.16	Variable noun class assignment in Maung	267
10.17	Classifiers and their interaction with other grammatical categories	269

11.1	Examples of physical properties in noun class assignment	277
11.2	Semantic basis of gender choice in German: an illustration	280
11.3	Noun classes in Bantu	282
11.4	Shape-based classes in ChiBemba	282
11.5	Size-based classes in ChiBemba	282
11.6	Noun classifiers for humans and deities in Jacalteco	284
11.7	Noun classifiers for non-humans in Jacalteco	285
11.8	Hole classifiers in Tzeltal	289
11.9	Function-based classifiers in Burmese	291
11.10	Verbal classifiers in Ojibway and Cree	297
11.11	Chipewyan (Athabaskan) classificatory verbs	298
11.12	Classificatory existential verbs in Kamoro (Asmat)	299
11.13	Preferred semantic parameters in classifiers	306
12.1	Burmese numeral classifiers for inanimate objects	312
12.2	Animate classifiers in Burmese	315
12.3	Verified superordinate-subordinate pairs in Japanese numeral classifiers	317
12.4	Reclassification of an inanimate noun in Burmese	319
13.1	Groups of nouns which tend to develop to classifiers	354
13.2	Numeral classifiers from body parts in Totonac	356
13.3	Noun classifiers derived from common nouns in Mam	357
13.4	Generic classifiers in Minangkabau	359
13.5	Semantic groups of verbs which develop into classifiers	362
13.6	A sample of verbal classifiers in Imonda	363
13.7	Historical changes in Mandarin Chinese classifiers	410
13.8	Typical sources for noun categorization devices	412
15.1	Scope of classifier types	427
15.2	Assignment of classifiers	429
15.3	Morphological realization of classifiers	430

List of Diagrams

2.1	Tendencies for animacy-based and shape-based noun classes	76
6.1	Verbal classifiers in Palikur	164
7.1	Locative classifiers in Palikur	173
7.2	Classifiers with spatial semantics in Toba	179
8.1	Semantics of numeral classifiers in Minangkabau	189
9.1	Semantics and form of demonstratives with classifiers in Tariana	239
9.2	Semantics and form of articles with classifiers in Tariana	240
10.1	Animacy hierarchy and expression of number	247
11.1	Gender assignment in Manambu	278
11.2	Extendedness in Proto-Bantu noun classes	283
11.3	Numeral classifiers in Totonac	289
11.4	Numeral classifiers for inanimates in Palikur	290
11.5	Interaction of semantic domains in numeral classifiers in Minangkabau	292
11.6	Classificatory verbs in Western Apache	298
11.7	Semantics of fourteen numeral classifiers in Akatek	302
11.8	Semantics of fourteen noun classifiers in Akatek	302
11.9	Polygrammaticalization of <i>batang</i> 'tree' in Minangkabau	302
11.10	Classifiers in Nambiquara	304
11.11	Semantics of classifiers in Kilivila	304
12.1	Structure of the <i>tua</i> category in Thai	314
12.2	Factors regulating the use of classifiers in Malay	324
12.3	Semantic network of the <i>nge</i> -class in Maasina Fulfulde	346
12.4	Gender pronouns in former and contemporary prescribed English usage	350
13.1	Evolution of gender markers (1)	367
13.2	Evolution of gender markers (2)	368
13.3	Phonological reduction in the development of noun classifiers in Mixtec	376
13.4	Singular and plural noun classes in Grebo	377
13.5	Gender in Proto-Dravidian (singular)	378
13.6	Gender in Proto-South-Dravidian (singular)	378

List of Abbreviations

A	subject of a transitive verb	DEF	definite
ABL	ablative	DEIC	deictic
ABS	absolutive	DEM	demonstrative
ACC	accusative	DER	derivational
ADJ	adjectivizer	DET	determiner
ADV	adverb	DIM	diminutive
AFF	affix	DIR	directional
AGR	agreement	DIST	distal
ALIM	alimentary	DS	different subject
	possession	DU	dual
AN, ANIM	animate	DUR	durative
ANA	anaphoric	EMPH	emphatic
ARG.MAN	argument manip- ulating derivation	ERG	ergative
		EXCL	exclusive
ART	article	EXT	extended
ASP	aspect	EYEW.PRES	eyewitness present
ATT	attributive		
AUG	augmented	F, FEM, f, fem	feminine
AUX	auxiliary	FRUST	frustrative
BENEFACT	benefactive	FUT	future
CAUS	causative	GEN	generic
CL	classifier	GN	genitive
CL:HAB	classifier: habitat	HAB	habitual
CMPL	completed	HON	honorific
COLL	collective	HORIZ	horizontal
COMIT	comitative	HUM	human
COMPL	completive	HUMIL	humiliative
CONC	concordial	IMAG	imaginary
CONJ	conjunctive	IMP	impersonal
CONT	continuous	IMPF	imperfective
CONTR	contrast	INAN, INANIM	inanimate
CURV	curved	INCL, incl	inclusive
CV	connective	INDEF	indefinite
CYLIDR	cylindrical	INS	instrumental
DAT	dative	INT	intensifier
DEC	declarative	INTER	interrogative
		IRREG	irregular form or shape

LIG	ligature	PRED	predicative
	vowel	PREF	prefix
LINK	linker	PRES	present
LIV.BEING	living	prim	primary
	being	PRO	1/2/3 person
LOC	locative		proform
M, MASC, m, masc	masculine	PROB	probability
MENS	mensural	PROGR	progressive
	classifier	pron	pronoun
MIN	minimal	PURP	purposive
MOD	modal	PX	proximity
N, NEUT	neuter	QUAL	qualifier
NCL	noun class	QUANT	quantifier
NEG	negative	RE	referential
NF, nf	non-	REC	reciprocal
	feminine	REFL	reflexive
NOM	nominative	REL	relativizer
NONPOSS, NPOSS	non-	REL.CL	relational
	possessed		classifier
NP	nonpast	REM.P.INFR	remote past
NUM.CL	numeral		inferred
	classifier	RES	resultative
O	object of a	S	subject of an
	transitive		intransitive verb
	verb	sec	secondary
OBJ	object	sg, SG	singular
PART	participle	SP	species
PASS	passive	SUBJ	subject
pcl	paucal or	SUBORD	subordinating
	plural	SUFF	suffix
	number	TA	tense-aspect
PERF	perfective		marker
pf	perfect	TAM	tense-aspect-
PI	past		mood marker
	imperfective	TH	thematic
PL	plural	THEM.CONTR	thematic contrast
POSS	possessive	TNS	tense
POSS.CL	possessive	TOP	topic
	classifier	TOP.ADV	topic advancing
PP	past		voice
	perfective	TOP.O	topical O
PRECONT	precontem-	VB	verbalizer
	porary tense	VCL	verbal classifier
		VERT	vertical

1 Preliminaries

1.1. General remarks

Almost all languages have some grammatical means for the linguistic categorization of nouns and nominals. The term ‘classifiers’ will be used here as an umbrella label for a wide range of noun categorization devices. Different types of classifier can be distinguished by their grammatical status, degree of grammaticalization, conditions for use, meaning, kinds of origin, mode of acquisition, and tendencies towards loss.

Classifiers and noun categorization devices have long been a particular focus of interest in functional typology. The urgent need to establish a comprehensive typology of classifiers is motivated by a number of factors. First, a large amount of new data on classifier systems has been produced during the past decades; on the one hand, this data needs to be systematized, and on the other hand, its existence creates the opportunity of providing a typology with reasonable scope and validity. Second, due to the lack of an overarching unified analysis of classifier systems in the languages of the world, there exists a pervasive terminological confusion in the literature which makes difficult the cross-linguistic comparison of noun categorization devices as well as the analysis of new data. This book is an attempt to provide such a comprehensive approach insofar as this is possible at our present stage of knowledge about the structure and mechanisms of human languages and human cognition. The book is also intended to serve as a guide for analytic work on previously undescribed languages and their mechanisms for noun categorization.

Examples of different kinds of classifier are provided in §1.2. In §1.3 I briefly describe the theoretical framework used in this study, together with the database and sources. The next section provides a short overview of previous approaches to noun categorization which are precursors to the approach adopted here. The methodological basis for this approach is outlined in §1.5. The structure of this book is outlined in §1.6.

1.2. Classifiers: an illustration

Classifiers come in different guises.

Some languages have grammatical agreement classes, based on such core semantic characteristics as animacy, sex, or humanness. These are called

2 *Classifiers*

NOUN CLASSES, or GENDERS. The number of noun classes varies—from two, as in Portuguese (examples below), to ten, as in Bantu, or even to several dozen, as in some South American languages. Examples 1.1 and 1.2, from Portuguese, illustrate masculine and feminine genders which are marked on the noun itself and on the accompanying article and adjective.

1.1. o menin-o bonit-o
ART:MASC.SG child-MASC.SG beautiful-MASC.SG
'the beautiful boy'

1.2. a menin-a bonit-a
ART:FEM.SG child-FEM.SG beautiful-FEM.SG
'the beautiful girl'

A classifier can just categorize the noun by itself, as in the following example from Yidiny, an Australian language (Dixon 1982: 192 ff.). This is a NOUN CLASSIFIER.

1.3. bama waguja
CL:PERSON man
'a man'

Other languages have special morphemes which only appear next to a numeral, or a quantifier. They may categorize the referent of a noun in terms of its animacy, shape, and other inherent properties. These are NUMERAL CLASSIFIERS. The way they are used is exemplified with a shopping list in Japanese (Rie Hasada 1995) given in Table 1.1.

TABLE 1.1. *Shopping list in Japanese*

Shopping list	Numeral	Classifier	Meaning of classifier
<i>nasu</i> (eggplant)	<i>nana</i> (7)	<i>-ko</i>	CL:SMALLEQUIDIMENSIONAL
<i>kyuuri</i> (cucumber)	<i>hachi</i> (8)	<i>-hon</i>	CL:ELONGATED
<i>hamu</i> (ham)	<i>juu</i> (10)	<i>-mai</i>	CL:SHEETLIKE

A special morpheme may characterize a possessed noun in a possessive construction, as in 1.4, from Tariana, a South American language from the Arawak family. This is a POSSESSED classifier.

1.4. tfinu nu-ite
dog 1SG-CL:ANIMATE
'my dog'

A special morpheme in a possessive construction may characterize the way in which the referent of a possessed noun relates to that of the possessor. This is illustrated in 1.5 and 1.6, from Fijian, an Austronesian

language (Lichtenberk 1983a: 157–8). Such morphemes, underlined in 1.5 and 1.6, are called **RELATIONAL CLASSIFIERS**.

- 1.5. na me-qu yaqona
 ART CL:DRINKABLE-my kava
 ‘my kava’ (which I intend to drink)

- 1.6. na no-qu yaqona
 ART CL:GENERAL-my kava
 ‘my kava’ (that I grew, or that I will sell)

VERBAL CLASSIFIERS appear on the verb, but they categorize a noun, which is typically in S (intransitive subject) or O (direct object) function, in terms of its shape, consistency, and animacy. Example 1.7, from Waris, a Papuan language (Brown 1981: 96), shows how the classifier *put-* ‘round objects’ is used with the verb ‘get’ to characterize its O argument, ‘coconut’.

- 1.7. sa ka-m put-ra-ho-o
 coconut 1SG-to VCL:ROUND-GET-BENEFAC-T-IMPERATIVE
 ‘Give me a coconut’ (lit. ‘coconut to-me round.one-give’)

There are two more, much rarer, kinds of classifiers. Those which occur on locative adpositions, are called **LOCATIVE CLASSIFIERS**. This is illustrated with 1.8 and 1.9, from Palikur, an Arawak language from Brazil.

- 1.8. pi-wan min
 2SG-arm ON+VERT
 ‘on your (vertical) arm’

- 1.9. ah peu
 tree on+BRANCH.LIKE
 ‘on (branch-like) tree’

Classifiers which are associated with deictics and articles are called **DEICTIC CLASSIFIERS**. Examples of deictic classifiers, from Mandan, a Siouan language (Barron and Serzisko 1982: 99), are given in 1.10 and 1.11.

- 1.10. de-māk
 ‘this one (lying)’
 1.11. de-nak
 ‘this one (sitting)’

The term ‘classifier systems’ is used to denote a continuum of methods of noun categorization. Well-known systems, such as the lexical numeral classifiers of Southeast Asia, on the one hand, and the highly grammaticalized gender agreement classes of Indo-European languages, on the other, are the extremes of this continuum. They can have a similar semantic

basis; and one type can develop out of the other. Parameters used for the proposed typology of classifiers are discussed in §1.5.

1.3. Theoretical framework, data, and sources

The aim of this book is to present a functional-typological, empirically based account of noun categorization devices across the languages of the world. The analysis is cast in terms of basic linguistic theory, 'the fundamental theoretical apparatus that underlies all work in describing languages and formulating universals about the nature of human language', where 'justification must be given for every piece of analysis, with a full train of argumentation' (Dixon 1997: 132; see also Dixon 1994: p. xvi). The categories, and their properties, considered here are developed inductively.¹

This study is based on examination of the grammars of about 500 languages representing each major language family and each linguistic area across the globe. A large database has been used, since the presence or absence of a particular kind of classifier system is often an inherited property of a language family or a diffusional property of a linguistic area. Special attention has been paid to data that has recently become available on the languages of South America (which by and large have not been included in previous typological studies of classifier systems). Data on the following languages come from my own fieldwork: Tariana, Baniwa, Warekena, Bare (Arawak family), Tucano, Piratapuya (East Tucano family), Paumari (Arawá family), from Brazil; and Manambu (Ndu family, East Sepik) from Papua New Guinea.

I have not restricted myself to considering just some samples of the available set of languages. Rather, I have looked at every language on which I could find data and which has noun categorization devices. This approach (sometimes called 'sample of convenience') allowed me to make the typology proposed here as comprehensive as it could be at our present level of knowledge about the languages of the world, without imposing artificial limitations dictated by this or that 'sampling strategy'. Owing to limitations of space, I could not cite all the examples of occurrence of every particular phenomenon. I usually provide a particularly illustrative example, and mention others. If a certain phenomenon is found in more than half of the languages under consideration I call it 'relatively frequent'; if it is found

¹ Cf. Bloomfield (1933: 20) 'The only useful generalizations about language are inductive generalizations. Features which we think ought to be universal may be absent from the very next language that becomes accessible. . . . The fact that some features are, at any rate, widespread, is worthy of notice and calls for an explanation; when we have adequate data about many languages, we shall have to return to the problem of general grammar and to explain these similarities and divergences, but this study, when it comes, will not be speculative but inductive.'

in a restricted number of languages (one to ten), I cite all of them and indicate its rarity. Note, however, that what appears rare to us at the present stage of knowledge may turn out to be frequent when we start learning more about hitherto little-known languages and areas. This is the reason why I choose not to give any statistical counts at this stage. Five hundred is no more than about one-tenth of all human languages, and it seems most judicious to follow a qualitative approach at the present time, postponing quantitative analysis until more data is available and can be assessed.

Lists of languages, of language families, and of linguistic areas considered, are given in the index. I chose not to enumerate classifier types found in each particular language referred to in the index in order not to impose my analytic solution onto a language which is not my area of expertise (readers can do this for themselves). Examples which come from my own work are not followed by the indication of a source. I preserve the orthography of the source (or use an accepted practical orthography, transcription, or transliteration) unless otherwise indicated.

1.4. Approaches to the typology of classifiers

Classifiers and noun categorization systems have long been a particular focus of interest in functional typology. They provide a unique insight into how people categorize the world through their language. The study of classifiers and noun categorization systems is intrinsically connected with many issues which are crucial in modern linguistics, such as agreement; processes in language development and obsolescence; the distinction between inflection and derivation; and types of possessive construction.

Noun classes and genders, on the one hand, and numeral classifiers, on the other, have been the object of linguistic investigation for as long as languages with these categories have been studied. The first overview full of fascinating insights—albeit preliminary—was provided by Royen (1929). A number of linguists have had ideas about similarities between different systems of noun categorization devices; for instance, Worsley (1954) pointed out functional similarities between Bantu-type noun class systems and noun classes and numeral and verbal classifiers in Anindilyakwa, an Australian language.

The systematic typological study of classifiers started only about two decades ago. Studies of classifiers divide into two kinds: attempts to create a general typological picture, and studies of individual types. The two cannot be easily separated, since each discovery of a new type provides feedback into the general typological picture.

During the last two decades, there have been a number of proposals for a semantic and grammatical typology of noun categorization systems (often also called 'noun classification'; e.g. Dixon 1968; 1982; Denny 1976; Allan 1977; Craig 1986a). Recently the typological parameters of classifiers and other agreement categories have had to be revised in the light of new data, especially those from previously undescribed South American Indian languages (e.g. Derbyshire and Payne 1990; Craig 1992, forthcoming; Corbett 1991).

Greenberg undertook a pioneering study of classifiers, in his paper on numeral classifiers and substantival number (1972). Though this paper does not overtly suggest any typology of noun categorization devices, various classificatory phenomena are mentioned alongside numeral classifiers (e.g. relational classifiers in Oceanic languages and verbal classifiers); he also suggested a correlation between the existence of numeral classifiers in a language and other grammatical categories, such as obligatory expression of number.

Further attempts at global typologies of classifiers include Adams and Conklin (1973), Denny (1976), Allan (1977), and Serzisko (1982). Dixon (1982) put forward an important suggestion for distinguishing between the two extremes of noun categorization devices: obligatory grammatical noun class systems, and semi-open lexical-like systems of classifiers (e.g. noun classifiers and numeral classifiers). Dixon (1982; 1986) was also the first to have explicitly stated a correlation between language type and noun categorization devices (that classifiers tend to be a property of isolating languages, while noun classes tend to be present in fusional and agglutinating languages); he showed how one type (noun classifiers) can develop into another (noun classes). The distinctions he drew between noun classes and classifiers are shown in Table 1.2.

TABLE 1.2. *Differences between noun classes and classifiers*

	Noun classes	Classifiers
Size	Small finite set	Large number
Realization	Closed grammatical system	Free forms
Scope	Marking is never entirely within the noun word	Never any reference outside the noun phrase

Source: Dixon (1982; 1986).

Allan (1977) provided a useful overview of noun categorization, for the first time explicitly stating that the following types of noun categorization device belong to the same domain: noun classes (or concordial classifiers), numeral classifiers, verbal classifiers (including separate mor-

phemes and suppletive classificatory verbs), possessive and intralocative classifiers.

Serzisko (1982) considered gender, noun class, and numeral categorization as a part of a continuum of 'classificatory techniques' (under the typological dimension of 'apprehension'), working out correlations between these and other categories (such as number), and comparing them as to their grammaticality, semantic complexity, and variability.

A very important though frequently underestimated contribution to the typology of noun categorization is found in Seiler and Stachowiak (1982) and Seiler and Lehmann (1982), followed by a summary in Seiler (1986), and also in Seiler's (1983) book on possession. These volumes are full of insightful case studies; also, Barron and Serzisko (1982), followed by a summary in Seiler (1986), provided the first consistent evidence in favour of the existence of deictic (or article) classifiers in Siouan languages. Seiler (1986) was the first to put forward the view of various kinds of classificatory techniques—including numeral classifiers, verbal classifiers, noun classes and 'article' classifiers—as continua within the broad dimension of apprehension.

Craig (1986a) was a major contribution to typological studies on noun categorization, their role in cognition and culture. In particular, noun classifiers as a special type have been established on the basis of her work. A new view on the typology of noun categorization devices was provided by Derbyshire and Payne (1990) in their survey of typologically unusual systems of noun categorization devices in Lowland Amazonian languages. Amazonian languages were shown to systematically allow more than one—and often more than two—types of noun categorization simultaneously.

Further typological studies on classifiers include Nichols (1989b), Kiyomi (1992), and Croft (1994). These focused on different parameters. Nichols concentrated on the morphosyntactic realization of classifiers, pointing out the differences between agreeing and non-agreeing noun categorization devices. Kiyomi (1992) attempted to establish morphosyntactic correlates of classifier realization (with free or with bound morphemes) for the main classifier types, and argued that neither animacy nor shape can be established as defining semantic parameters for a typology of noun categorization devices. See Table 1.3.

TABLE 1.3. *Classifiers, their morphological realization, and semantics*

Free morpheme classifiers	Bound morpheme classifiers
Nominal classifiers (Animate, Shape)	Concordial classifiers (Animate, Shape)
Non-nominal classifiers (Animate only)	Predicate classifiers (Animate, Shape)
	Intralocative classifiers (Shape only)

Source: Kiyomi (1992: 33).

Croft (1994) reanalysed classifier types, associating each of them with semantic and pragmatic functions (he disregarded a few problematic classifier types, such as locative and deictic classifiers, and systems of verbal classifiers with no animacy distinctions; see Chapters 6, 7, and 11 below). See Table 1.4.

TABLE 1.4. *Classifiers and their functions*

Classifier type	Semantic/pragmatic function
Noun class	Determination (reference)
Numeral classifiers	Enumeration
Possessive classifiers	Possession
Predicate classifiers	Spatial predication

Source: Croft (1994: 147).

Recent overviews of the typology of classifier systems can be found in an in-depth study of Japanese numeral classifiers by Downing (1996), and in a detailed analysis of classifiers in Kilivila (Austronesian) by Senft (1996).

Craig (1992; forthcoming) argued for the existence of the following types of classifiers based primarily on the morphosyntactic loci in which they occur: numeral classifiers; noun classifiers; noun class and gender; verbal classifiers; genitive classifiers. Craig (forthcoming: 42) also mentioned the existence of a 'marginal' classifier type—classifiers which occur with articles or deictics. Further arguments in favour of this morphosyntactic typology include cooccurrence of types within one language, different semantics for distinct classifier types, and different degrees of grammaticalization of classifiers. Importantly, classifiers are not presented as discrete types, but rather as focal points on various continua. This prototype-continuum approach, which implies a gradient rather than categorical treatment of properties of classifier systems, is taken up in the present study (see §1.5).

Craig's approach was elaborated upon in a case study of classifiers in Tariana by Aikhenvald (1994a), and in Palikur by Aikhenvald and Green (1998). The typology proposed in this book is largely based on the schema established by Craig.

However, the current literature is somewhat confusing as far as generally adopted definitions and concepts are concerned. The way linguists of different traditions and theoretical trends use different terms, such as GENDER, NOUN CLASS, CLASSIFIER, can be misleading.

The terms GENDER and NOUN CLASS are sometimes used interchangeably (see §2.1). Corbett (1991) uses 'gender' as a cover term for agreement classes, while Evans (1997: 109) opts for 'noun class' to cover the same

phenomenon. GENDER has also been used in a quite different way. In the Athabaskan linguistic tradition the term 'gender' is used to refer to verbal classifiers which mark agreement with intransitive subject or transitive object, and characterize the referent noun in terms of shape and form (Thompson 1993). In the Bantuist tradition, the term 'noun class' is used to refer to a set of singular and of corresponding plural forms of a noun and the agreement markers they trigger on modifiers and on the predicate, while the pairs of singular and plural markers are considered 'genders'. For instance, Singular Noun Class 1 forms one 'gender' with its plural counterpart, Noun Class 2 (see e.g. Table 2.1).

The term 'verbal classifier' is sometimes used by Australianists (Silverstein 1986; Green 1989; Reid 1990 and p.c.; Rumsey 1982; Donaldson 1980 and others) to refer to a closed class of inflected verbs which typically carry grammatical marking, and 'classify' the lexical verb by delimiting its aspect or scope (e.g. 'do something on the surface', 'do something with hands', 'do moving up'). There is typically a small class of inflected verbs with fairly generic meanings (often called 'simple' verbs, e.g. Rumsey 1982, for Ungarinjin; Silverstein 1986) which together with a 'main verb' (or 'co-verb') form a complex verb. In Ngiyambaa (Donaldson 1980: 201-24; Dixon forthcoming: §6.2) a main verb 'dig', 'sew', or 'spear' takes the classifier 'pierce'; and a verb such as 'take', or 'pick up' requires a classifier 'do with hands'. This usage of 'classifier' has some similarity with noun categorization via generic noun classifiers: a simple, or 'classifier' verb defines the generic scope of action, and the main verb specifies it; similarly, a noun classifier indicates general reference (e.g. 'person' for people or 'animal' for animates), and the specific noun following it further specifies this reference. This usage is completely different from the one adopted here; however, as pointed out by Ian Green and Nicholas Reid (p.c.), simple verbs may develop further semantic specifications whereby they start being used to characterize the particular kind of instrument or location. Further study is needed to delineate 'noun classifying' functions of simple verbs in Australian languages.

The term 'verbal classifiers' is used in another, completely different way by some specialists in the languages of South and Southeast Asia. Haas (1942: 205) calls 'words indicating how many times an event takes place' verbal classifiers; i.e. in a sentence like 'he ran twice' 'twice' is considered a verbal classifier. This term is employed in a similar way for Newari (Tibeto-Burman) by Bhaskararao and Joshi (1985: 17), and for Mulao, a Tai language, by Jun and Guoqiao (1993: 48).

The term 'classifier' is used in yet another way in the Athabaskan linguistic tradition, where it refers to markers of voice and change of transitivity which have nothing to do with categorization of nouns.

Some authors simply avoid the term 'classifier'. Moussay (1981) uses the

term 'spécificatif' for numeral classifiers and 'catégoriel' for noun classifiers in Minangkabau.

A number of the general statements about different types of classifier have recently been shown to be erroneous. Some of the previously accepted universals and general tendencies do not, in fact, hold. For example, Dixon (1982: 220) suggested that languages could not have classifiers and gender as separate categories, and stated: 'no example is known of a language with two distinct systems of noun classes' (see also Craig 1986a; 1986b; 1986c). Recent work on South American and Papuan languages has shown that classifiers and genders do cooccur, and that languages can have two distinct systems of noun classes. For instance, Baniwa, a North Arawak language from Brazil, has a system of two genders, and also a system of over 40 noun classes (see Chapters 8 and 9 below).

The dichotomy between a concordial noun class as an 'obligatory grammatical system where each noun chooses one from a small number of possibilities' and noun categorization as a system where 'noun classifiers are always separate lexemes which may be included with a noun in certain syntactic environment' (Dixon 1986: 105) appears to be rather simplistic, especially in the light of the data from Amazonian languages. The presence of noun classes had often been associated with a fusional or agglutinating morphological type, and classifiers (especially numeral classifiers) were viewed as a typical property of isolating languages—a premise that also appears to be a little simplistic when viewed cross-linguistically.

Finally, particular terms, such as 'classifier', 'noun classifier', or 'noun categorization system', are frequently used by different authors either in a different way for different types of system or as a cover term for any kind of system. Thus, it is not always clear what is a classifier and what is a concordial noun class in each particular case.

During the last two decades, a number of studies of specific classifier types and individual languages have made an important contribution to an overall typological picture. Corbett's (1991) book on GENDER (which is used as a cover term for NOUN CLASS SYSTEMS) is an important, almost encyclopedic, overview of this type of noun categorization. It is almost impossible to enumerate all the studies of noun class systems in African languages; however, the collection *La Classification nominale dans les langues négro-africaines* (1967) and Hyman's (1980) book on noun classes in Grasslands Bantu languages remain the main reference on the subject. Heine's (1982a) article 'African Noun Class Systems' remains the main reference for the typology of noun classes in African languages. Noun class systems in Papuan languages are described by Foley (1986); some of these are extremely unusual—their assignment may be largely based on phonological form (see Foley 1986; 1991; Conrad 1996; Nekitel forthcoming). Work on noun classes in Australian languages includes ground-breaking studies by

Dixon (1972; 1982), Sands (1995), and the papers in Harvey and Reid (1997).

NOUN CLASSIFIERS have been introduced into the typological picture comparatively recently; their properties have been discussed at length by Dixon (1982), Craig (1986a; 1992; forthcoming), and in more specific case studies, e.g. Zavala (1992; forthcoming), Reid (1997), and Sands (1995). In addition, Payne (1990) and Derbyshire and Payne (1990) considered the problem of noun classifiers in Amazonian languages.

There is an immense corpus of literature on NUMERAL CLASSIFIERS, especially in Southeast and East Asian languages. Adams (1989) and Downing (1996) provide in-depth discussions of problems relevant for the cross-linguistic definition of numeral classifiers (further, more language-specific, or area-specific case studies include Barz and Diller 1985; Goral 1978; Bisang 1993; 1996; Pe 1965; T'sou 1976; Conklin 1981).

The existence of CLASSIFIERS IN POSSESSIVE CONSTRUCTIONS in Oceanic languages was first recognized by Codrington (1885). The credit for the first systematic study of relational classifiers and how they differ from numeral classifiers in Oceanic languages goes to Lichtenberk (1983a); among further studies one must mention Harrison (1976) for Mokilese; Dixon (1988) for Fijian; Pawley and Sayaba (1990) for Wayan, a Western Fijian dialect; and Rehag (1981) and Keating (1997) for Ponapean. Seiler (1983) provides an insightful analysis of noun categorization in possessive constructions and of the differences which can be noted between classification devices which characterize the ways in which nouns can be possessed, or handled (relational classifiers) and devices which describe properties of possessed nouns (possessed classifiers, in our terminology). Carlson and Payne (1989) attempted a broader survey of relational classifiers in some North American Indian languages (Yuman, Uto-Aztecán) and some South American Indian languages (some Carib, Tupí-Guaraní, and Jê languages); further data on relational and possessive classifiers in South American languages can be found in Rodrigues (1997, on Kipeá, an extinct language of the Karirí family, Macro-Jê, South America), Rodrigues (1999), Barnes (1990), Martins (1994), and Aikhenvald (1994a).

VERBAL CLASSIFIERS and SUPPLETIVE CLASSIFICATORY VERBS have been the subject of extensive study based on the facts of specific language families. Seminal studies of classificatory verbs in Athabaskan languages include Hoijer (1945), Davidson *et al.* (1963), Krauss (1968), Basso (1968), Carter (1976), and Thompson (1993); also see Mithun (1986) and Seiler (1986). There is extensive literature on classificatory verbs in other North American Indian languages, e.g. Kiowa-Tanoan (Speirs 1974) and Cherokee (Blankenship 1996; 1997). Verbal classifiers in South American languages are discussed in Derbyshire and Payne (1990); verbal classifiers in Papuan languages are considered by Lang (1975), Brown (1981), and Merlan *et al.*

(1997). For verbal classifiers in Mesoamerican and South American Indian languages see also Suárez (1983), Gonçalves (1987), and Mithun (1984; 1986).

A few studies have been undertaken on rare and problematic classifier types. The existence of DEICTIC classifiers as a special type has been shown by Klein (1979), Vidal (1995; 1997), and Céria and Sândalo (1995), for the languages of the Guaiacuran family of Argentina and Brazil. Barron and Serzisko (1982) describe article classifiers for Siouan languages. Further data from South American languages in support of the existence of deictic classifiers are given by Aikhenvald (1994a; forthcoming b). The existence of a special type of LOCATIVE classifier was first suggested by Allan (1977) (the term he used was 'intra locative'); his results were criticized by Croft (1994). In fact, locative classifiers have only been found in a limited number of South American languages (to which Croft did not have access), e.g. Palikur (Arawak), Dâw (Makú), and Carib languages (see Aikhenvald 1994a; 1996b).

Up until now no systematic attempt has been made to consider multiple classifier systems in a cross-linguistic perspective (see Chapter 8 below). Previous studies have not taken account of the unusual types of multiple classifier system found in South American Indian languages. (Systems of this kind were only briefly mentioned by Dixon 1982; Craig 1992; forthcoming; and Lichtenberk 1983a.) Among descriptions of multiple classifier systems from other parts of the world one should mention Hurd (1977) on the Nasioi language from Bougainville, and Worsley (1954) on Anindilyakwa, an Australian language from Groote Eylandt. Recently, the number of studies of multiple classifier languages has increased, e.g. Gonçalves (1987), on Mundurukú, a Tupí language from Amazonia; Bisang (1993) on Hmong, a Miao-Yao language from China; Onishi (1994) on Motuna, a Papuan language; Foris (forthcoming) on Sochiapan Chinantec from Mexico; Vidal (1997) on Pilagá, a Guaiacuran language from Argentina; Shepard (1997) on Machiguenga, an Arawak language from Peru; and the survey of multiple classifier systems in Arawak languages in Aikhenvald (1996b). Another problem for multiple classifier systems is 'fuzzy' boundaries between types which makes it difficult to attribute a given language to a particular type (see e.g. the discussion in Vidal 1997).

To summarize—in spite of the considerable work already accomplished, a new, integrated typological framework is needed to account for all the types of noun categorization device and the new language data which have appeared on the linguistic scene during the last decades. This is attempted in the present volume.

1.5. Parameters for the typology of classifiers

All human languages have some ways of categorizing nouns and their referents in terms of their semantic and syntactic properties. The purpose of this book is to investigate how languages employ classifiers to provide a semantically based categorization, which may have far-reaching implications concerning human cognitive mechanisms.

Classifiers are defined as morphemes which occur 'in surface structures under specifiable conditions', denote 'some salient perceived or imputed characteristics of the entity to which an associated noun refers' (Allan 1977: 285), and are restricted to particular construction types known as 'classifier constructions'. Classifier constructions are understood as morphosyntactic units (which may be noun phrases of different kinds, verb phrases, or clauses) which require the presence of a particular kind of a morpheme, the choice of which is dictated by the semantic characteristics of the referent of the head of a noun phrase.

Nouns and their referents can also be categorized in various other ways, e.g. by choosing different number forms for nouns with different semantics; by assigning the nouns to different declension classes; or by using different pronominalization strategies. These strategies of noun categorization (sometimes also called 'noun classification') are not considered classifiers. However, they may be used in a way functionally similar to classifiers, and they often reflect comparable semantic parameters. Historically, they may go back to classifier systems. Examples are given in Appendix 1.

The main purpose of this book is to present a typology of classifiers primarily based on the morphosyntactic loci (or environments) of classifier morphemes (following the approach in Craig 1992; forthcoming). This implies establishing types of noun categorization system which acquire surface realization in natural languages. As a result, the typology is inclusive in that it covers types of classifier morpheme and construction types in which they are required, and categorization types. We start with a typology of classifier morphemes and the constructions in which they are employed, and then proceed to uncover a link between these and universal and language specific parameters of categorization types. This is the basis for distinguishing definitional properties and contingent characteristics of classifier types.

The terminology chosen for each classifier type relies as much as possible on currently accepted terminology. If there are several terms in use, I employ the one which is most current and most transparently describes the morphosyntactic locus of a classifier type (e.g. I use 'verbal classifier' rather than 'verb-incorporated classifier').

Following Craig (forthcoming: 43), classifier types are not viewed as discrete entities, but rather as focal points on continua of various properties used for the present typology (see below). As the result, definitional as well

as secondary, or contingent, properties of different classifier types will be shown to be gradient rather than categorical; this accounts for the existence of instances of classifier systems which 'do not fit squarely into any of the types' (Craig forthcoming: 43). As Frawley (1992: 30) puts it,

if we look at ordinary language, we find that it is full of gradient phenomena, more technically known as *fuzziness*. . . . The insight behind fuzziness indicates that categories have vague boundaries and are internally organised from central focal values, the *prototype* (Rosch 1973, 1975a, b), to less focal instances and fringe values. As the centrality of the category fades, . . . criteria for membership in the category are less decisively applied, and categories merge into each other.

Consequently, classifier types outlined and argued for in this study correspond to *prototypes*, or *focal instances*, which display all the definitional and most of the contingent properties of a type. Less focal instances represent various points on *continua* for different parameters of a typology of noun categorization; these display varying degrees of the prototypical properties of each type. In describing and analysing the data on noun categorization devices in a given language, it is important to situate them within the continua of various gradient properties rather than to try and fit them into the mould of cross-linguistically established 'types'.

This prototype-continuum approach is also justified by historical facts about classifier systems—it is well known that distinct classifier types 'blend into one another through time' (Craig forthcoming: 43). These points will be amply illustrated within the present study; they are summarized in Chapter 15.

The following dimensions will be employed to establish focal points on the typological continuum of noun categorization devices.

(A) Morphosyntactic locus of coding

A noun categorization device can be realized in different morphosyntactic loci, that is, on the head, or on all—or just some—of the dependents. We will pay particular attention to languages which use different sets of classifier morphemes (often with different semantic and other properties) in several morphosyntactic environments. The coexistence of these sets in one language constitutes a strong argument in favour of the proposed typology, since this indicates the independent existence and independent development of different noun categorization devices in several morphosyntactic environments in one language.

Some kinds of noun categorization device have several distinct subtypes coexisting within one language: one set of noun classes may be used in one environment, and a somewhat different set in another. For instance, many Arawak languages of South America have a small system based on the masculine/feminine distinction realized on verbal cross-referencing markers

and on demonstratives, while adjectival modifiers show a large system of agreement noun classes. Systems of this kind are called 'split' systems; they may represent potential new 'focal points' for developing further classifier types.

(B) Scope, or domain of categorization

Noun categorization devices can refer to nouns within noun phrases of different structures (modifier-head, possessive noun phrases, or adpositional noun phrases), or within a verb phrase. They can also refer to different constituents (e.g. possessed noun or possessor; A, S, O, or an oblique argument). Thus, one can say that in 1.7 it is the O constituent, 'coconut', that is being categorized by the morpheme *put* 'classifier: round', and thus it constitutes the scope, or domain, of this classifier morpheme.

(C) Principles of choice, or 'assignment' of noun categorization devices

The choice of a classifier may depend on some semantic properties of the referent of the noun they categorize. However, it can also depend on other properties of a noun (e.g. morphological or phonological).

(D) Kind of surface realization

Some noun categorization devices are realized with an affix or a clitic, while others often appear as separate words.

(E) Agreement

Some noun categorization devices involve agreement, and some do not. Agreement is understood as a requirement in covariance between grammatical meanings of grammatical morphemes (cf. Steele 1978: 610; Lehmann 1982: 203; see §2.4 below). Categories which involve agreement are 'syntactic' (or 'inflectional') in nature.

(F) Markedness Relations

Some noun categorization devices have a functionally and/or a formally unmarked term; while others tend not to.

(G) Degree of Grammaticalization and Lexicalization

Some noun categorization devices are highly grammaticalized closed sets while others tend to involve a lexical choice. A more lexical kind of noun categorization can become grammaticalized.

(H) Interaction with other grammatical categories

Different types of noun categorization device tend to show different dependencies with other grammatical categories (such as number, or case, or verbal categories).

(I) Semantic organization of the system

Noun categorization in the languages of the world is based on a number of universal parameters (e.g. 'human' versus 'non-human'). However, noun categorization devices differ in terms of a number of other parameters, termed their 'preferred semantics'. They also differ as to the organization of their systems: in some, but not in others, every noun has to be assigned a classifier. They also differ in the degree of their semantic transparency and in the syntactic and discourse-pragmatic functions they perform. Classifiers of different types differ in how they respond to socio-cultural influence.

(J) Evolution and decay

Distinct types of noun categorization devices differ in their etymological sources, and in the ways they develop and how they fall out of use. Classifiers of one type can develop into another.

(K) Language Acquisition and Dissolution

Distinct noun categorization devices show fundamental differences in how they are acquired by children, and what processes they undergo under language dissolution in aphasia.

Properties (A–G) are definitional properties of classifiers, in agreement with the morphosyntax-prior approach to classifiers adopted here. Properties (H–K) are contingent properties. Once the types of classifiers are established with respect to characteristics (A–G), they will be shown to display correlations with properties (H–K).

Quite a few languages use different sets of morphemes in different classifier environments. Many languages employ the same (or almost the same) set of classifier morphemes in different morphosyntactic loci. In this case, the question to ask is whether we should consider them as instances of distinct, albeit homophonous, classifier types, or as basically one type extended to other environments. These and other related issues will be discussed together with the problems of multiple classifier systems.

The structure of this book, as outlined in the next section, follows the above order: we discuss the definitional properties of classifiers first, and then proceed to consider the contingent ones.

1.6. The structure of this book

We will first discuss the proposed types, or 'focal points' on the continuum of noun categorization devices with respect to their definitional properties (A–G above) in the following order.

NOUN CLASSES and **GENDERS** are noun categorization devices realized outside the noun itself within a head-modifier noun phrase. They are realized, as agreement markers, on modifiers such as adjectives, but may also appear on modifiers from closed classes such as demonstratives and interrogatives. They can also be realized outside the noun phrase, e.g. be marked on the predicate, or even on adverbs. They are most often affixes. They usually contain reference to inherent properties of nouns, such as animacy and sex, and sometimes also shape, structure etc. Some languages have a special smallish set of noun classes/genders restricted to closed classes of modifiers (demonstratives, and others) along with a different set which appears on modifiers from other classes. These are discussed in Chapter 2.

NOUN CLASSIFIERS are associated with the noun itself, and are independent of any other element in an NP, or in a clause. They may be independent words, or, more rarely, affixes attached to nouns. They refer to inherent properties of nouns. Noun classifiers are free forms. Noun classes and noun classifiers differ in their synchronic properties; however, noun classes often develop from noun classifiers. These are discussed in Chapter 3.

NUMERAL CLASSIFIERS are another kind of noun categorization device which operate within an attributive NP. These are realized outside the noun in a numeral NP, and/or in expressions of quantity. Numeral classifiers can be free forms, or affixes, typically to the numeral or quantifier. They refer to the noun in terms of its inherent properties. These are discussed in Chapter 4.

Noun categorization devices which operate within a possessive NP are considered in Chapter 5. They can be of three kinds:

- (i) The scope of categorization is the possessive relation itself, i.e. the way a noun can be possessed, or treated. These markers are called **RELATIONAL CLASSIFIERS**; they refer to the function of a noun, and not to its inherent properties.
- (ii) The scope of categorization is the possessed noun itself. Classifiers which categorize the possessed noun are called **POSSESSED CLASSIFIERS**. The noun is categorized in terms of its inherent properties.
- (iii) The scope of categorization is the possessor, and its inherent properties. These are **POSSESSOR CLASSIFIERS**.

Another type of classifiers which have a clause as their scope are **VERBAL (OR VERB-INCORPORATED) CLASSIFIERS** discussed in Chapter 6. Their scope is an argument of the predicate, usually in S/O function, more rarely in an oblique function, and they are realized on the verb. They refer to inherent properties of the noun; and may also convey information on its position in space.

There are a few further, rare and rather problematic kinds of noun categorization devices with an NP as their scope. LOCATIVE CLASSIFIERS appear in adpositional NPs attached to an adposition, and characterize the head noun in terms of its inherent properties. Some languages have DEICTIC classifiers—morphemes which appear on deictics within an NP and qualify the noun in terms of its inherent properties and its orientation, such as horizontal or vertical. These are considered in Chapter 7.

Some languages have more than one kind of noun categorization—these are discussed in Chapter 8. The same set of morphemes can be used in several classifier environments—see discussion in Chapter 9.

We then consider contingent properties of classifiers. The ways in which different classifier types interact with other grammatical categories are discussed in Chapter 10. Parameters for the semantic categorization of referents of nouns and the preferred semantics of different classifiers are considered in Chapter 11.

The semantic organization of classifier systems and their functions are dealt with in Chapter 12, together with a discussion of socio-cultural parameters and mechanisms of human cognition reflected in noun categorization. This chapter demonstrates the unitary basis for noun categorization devices, providing support for considering them as variant realizations of one phenomenon.

The origins, evolution, and decay of different noun categorization devices are discussed in Chapter 13. The processes noun categorization devices undergo in language acquisition and dissolution are considered in Chapter 14. The results of the proposed typology and perspectives for further studies are given in the concluding Chapter 15.

Appendix 1 describes noun categorization by means other than classifiers, i.e. through marking number, grammatical relations, and other categories. Appendix 2 contains additional examples of semantic changes in the process of development from nouns to classifiers.

Suggestions for linguists undertaking fieldwork on classifier languages are provided in Appendix 3.

2 Noun Class and Gender Systems

2.1. General remarks

NOUN CLASSES and GENDERS are grammaticalized agreement systems which correlate—at least in part—with certain semantic characteristics (particularly in the domain of human and animate referents). They are sometimes called concordial classes; they include grammaticalized ‘gender’ systems of the Indo-European type. They are realized outside the noun itself, usually on modifiers which most often include adjectives, but may also include modifiers from closed classes (demonstratives, interrogatives, possessives, etc). They can also be realized outside the noun phrase, i.e. be marked on the predicate, or even on adverbs. Some languages have a special smallish set of noun classes/genders restricted to closed classes of modifiers (demonstratives, and others) along with a different set which appears on modifiers from other classes.

A terminological clarification is in order. The term GENDER was first used in the 5th century BC by the Greek philosopher Protagoras, when he divided Greek nouns into three classes: ‘feminine’, ‘masculine’, and ‘inanimate’ (nowadays called ‘neuter’). This is a typical gender system, which is found in many Indo-European languages. Latin had a similar system; later, neuter nouns were redistributed between the other two genders, giving the modern system of masculine and feminine in French and Italian.

When Europeans came to study African languages, they discovered larger gender-like systems with eight or more distinctions in languages like Swahili; these often did not include a masculine/feminine distinction. The term NOUN CLASS came to be used for systems of this type.

NOUN CLASS, GENDER, and sometimes GENDER CLASS are often used interchangeably, depending on the linguistic tradition (some examples are given in §1.4). Here I shall use ‘noun class’ as a cover term for noun class and gender. In agreement with the linguistic tradition, I shall reserve the term gender for small systems of two to three distinctions (always including masculine and feminine), like the ones typically found in Indo-European, Afroasiatic, and Dravidian languages.

Since gender systems show some correlation with sex, many non-linguists (and a few linguists) erroneously confuse ‘linguistic’ gender and sex. However, sex represents biological categorization, and gender represents grammatical categorization. Feminine and masculine genders often

include inanimate nouns with no connection to female or male sex, e.g. French *maison* 'house' (feminine), *château* 'castle' (masculine).

The languages of the world differ in the number of noun classes they have, how much semantic transparency there is to noun class assignment, where and how noun class gets expressed, and whether it is possible to change the noun class of a given noun.

Noun class systems are typically found in languages with a fusional or agglutinating (not an isolating) profile. Noun class agreement is often a major criterion for distinguishing nouns from other word classes. In a language where noun and adjective have similar morphology, an adjective can usually take any noun class marking whereas a noun is normally restricted to one class.

Because of the limitations of space, it is impossible to cover all the literature on noun class systems. To avoid an overlap with Corbett's (1991) study of noun class systems (for which he uses a cover term 'gender'), I will concentrate on the issues and examples which have not been considered there, and briefly mention the ones for which Corbett (1991) provides detailed coverage.

The properties of noun class systems are considered in §2.2. Noun class assignment is discussed in §2.3. Noun class agreement is dealt with in §2.4. The next section discusses markedness relations and resolution in noun class systems. Realization of noun classes is analysed in §2.6. Some languages have two noun class systems. These may be in a complementary distribution with respect to the modifiers with which they are used; they may display peculiar agreement properties—see §2.7. Finally, §2.8 surveys the distribution of noun classes in the languages of the world.

2.2. Properties of noun class systems

A noun class system is the most grammatical means a language can use for the semantic categorization of nouns. As we shall see later, other noun categorization mechanisms are more lexical and often more semantically based.

Noun class systems have the following definitional properties.

1. Some constituent outside the noun itself must agree in noun class with a noun. Agreement can be with other words in the noun phrase (adjectives, numbers, demonstratives, articles etc.) and/or with the predicate of the clause or with an adverb. That is, noun class can be realized in a number of morphosyntactic loci (depending on the agreement rules in the language) and its scope can be a noun phrase and/or a clause. Noun classes are defined syntactically. They constitute a closed obligatory grammatical

system (which often arises as the result of grammaticalization of some other noun categorization device: see Chapter 13).

Noun classes are realized with affixes or with clitics, and in most cases there is a limited, countable number of noun classes (see §2.6).

2. Noun class membership is assigned on semantic—and sometimes also morphological and phonological—principles. Each noun in the language belongs to one (or occasionally more than one) class(es).

There is always some semantic basis to the grouping of nouns into classes, but languages vary in how much semantic transparency there is. This semantic basis usually includes animacy, humanness and sex, and sometimes also shape and size. We will return to this in Chapter 11.

In some languages, in addition to the realization of noun classes through agreement, there is a marker of noun class on the noun itself, or on some nouns; in other languages nouns bear no overt marker.

Languages often have portmanteau morphemes combining information about noun class with number, person, case, etc. This is considered in Chapter 10.

Some systems based on animacy and sex (and traditionally called ‘gender systems’) do not, in fact, satisfy the criteria set out here. English distinguishes three genders just in 3rd person pronouns, *he/she/it*. They involve the opposition: male, female, inanimate. There are a few conventionalized metaphorical extensions, e.g. ships are commonly referred to with the feminine pronoun *she* (see further examples in §12.3.3). There is no gender agreement within a noun phrase or with a verb in a clause.¹ Gender markers in English simply have an anaphoric function, as they also do in Japanese where masculine and feminine forms are distinguished only in 3rd person pronouns with a human referent *kare* ‘he’, *kanojo* ‘she’ (recently introduced, possibly, under the influence of European languages: Walter Bisang, p.c.).² Many languages of the world also have animacy-based distinctions in interrogative and indefinite pronouns (see Haspelmath 1997), e.g. English *anybody* or *anything* which can be used anaphorically. Strictly speaking, these are not noun classes.

The presence of agreement is the main definitional property of a noun class. Some languages have singular/plural alternations which can be shown as at least partially conditioned by the semantics of nouns. However, these pairings do not correspond to different agreement classes. Such appears to be the case with singular/plural alternations in Eastern Sudanic

¹ However, if we follow Lehmann’s (1982: 219) view of agreement, the use of these pronouns to ‘agree’ in animacy/sex with their antecedent can be considered ‘anaphoric agreement’; see §2.4.1.

² There can also be complicated relations between sex of speaker and form of other pronouns, e.g. first person; these relate to the category of politeness: cf. Chapter 10.

languages (Dimmendaal forthcoming). This is a system of classifying nouns and their referents; however, it cannot be considered a system of noun classes. In contrast, Bantu languages have large systems of noun class affixes which are portmanteau morphemes of noun class with number; since they appear both on the noun itself and on the agreeing constituents they 'qualify' as noun classes. Modern Hebrew distinguishes two genders, masculine and feminine, both in the singular and in the plural (see §2.3.4) which are realized in agreement within a noun phrase and on the verb. Nouns also fall into several classes depending on their number and case forms (Aikhenvald 1990: 48); this second kind of classification lies outside the scope of the present study.

2.3. Principles of noun class assignment

The principles by which nouns are 'assigned' to different classes can be governed by semantics (§2.3.1), or formal morphological (§2.3.2) or phonological (§2.3.3) properties of a noun, or a combination of these (§2.3.4) (also see Corbett 1991: 7–69). In a sense all systems of noun class assignment are mixed, since there is always a semantic core which involves the universal semantic parameters (see §11.2.1) of sex, humanness, animacy but this is never the entire story.

Noun class systems were defined above as obligatory grammatical systems, such that every noun has to belong to a noun class. However, noun class assignment is sometimes impossible for a smallish group of nouns. These are exceptions to the statement that every noun in a language with a noun class system has to be assigned to a noun class. For example, Russian does not distinguish genders in the plural. Then, the gender of pluralia tantum, i.e. nouns which are used only in the plural and always have plural agreement, cannot be determined, e.g. *sani* 'sledge', *brjuki* 'trousers', *seni* 'entrance into a hut'.³

The psycholinguistic reality of gender assignment has been confirmed by recent studies of child language acquisition (Connelly 1984; Mills 1986; Tsonope 1988). We shall return to this in Chapter 14.

2.3.1. Semantic assignment

In languages with purely semantic assignment the class of a noun can be inferred from its meaning. In Tamil (Dravidian) all nouns divide into what are traditionally called 'rational' and 'non-rational' classes. 'Rational' nouns comprise humans, gods, and demons (Asher 1985: 136). Other

³ These nouns do distinguish animacy; see (C) in §2.4.4 on subgenders in Russian.

Dravidian languages, Malto, Kolami, Ollari, and Parji distinguish male humans as distinct from other nouns which refer to 'rational' beings. From the Northeast Caucasian family, Godoberi, Akhvakh, and Bagval distinguish male 'rational', female 'rational', and the rest (Corbett 1994b). Diyari (Australian), Kaingang (Jê), and the North Arawak subgroup divide nouns into female humans and the rest.

Semantic assignment can be more complex. Dyirbal (Dixon 1972: 306–12) has four classes. Three are associated with one or more basic concepts: gender 1—male humans, non-human animates; gender 2—female humans, water, fire, fighting; gender 3—non-flesh food. Gender 4 is a residue class, covering everything else. There are also two rules for 'transferring' gender membership. By the first, an object can be assigned to a gender by its mythological association rather than by its actual semantics. Birds are classed as feminine by mythological association since women's souls are believed to enter birds after death. The second transfer rule is that if a subset of a certain group of objects has a particular important property, e.g., being dangerous, it can be assigned to a different class from the other nouns in that group. Most trees without edible parts belong to gender 4, but stinging trees are placed in gender 2.

Mythological association plays an important role in class assignment in other languages, too. In the Western Torres Strait language all nouns denoting males are masculine, with the remainder being feminine. However, the moon is masculine, due to its mythological association with masculinity (Bani 1987). This is also characteristic of other Australian languages. The assignment of masculine and feminine noun classes in Abu' Arapesh (Papuan; Nekitel 1985, 1986, forthcoming) can often be explained by mythological associations; for instance, cassowary is feminine because it used to be a mythological woman, and the moon is masculine—it is a mythological man who engages in a sexual intercourse with women (making them menstruate).

In Ket (Krejnovič 1961, Dul'son 1968: 62ff.), all sex-differentiable nouns are masculine or feminine. Among non-sex-differentiable nouns, those which show a higher degree of activity or are particularly important for Ket culture are masculine, e.g. wood, large wooden objects, growing trees. Gender assignment of sun (feminine) and moon (masculine) is determined by their role in myths.⁴ Other inanimate nouns are treated as neuter.

The degree of semantic motivation for noun classes varies from language to language. Systems with a larger number of noun classes tend to have

⁴ Also see Harvey (1997) for an insightful overview of semantic parameters employed in gender assignment in Australian languages. In this and other cases it can also be argued that the mythological role was originally determined by gender assignment, and not the other way round. Indeed, in many cases it is impossible to prove which line of argument is better founded.

more semantic motivation; however, this is not necessarily so. Languages of the Ndu family (East Sepik, Papua New Guinea) have two semantically assigned genders (see §2.4.3). In Babungo (Grassfields Bantu, Benue-Congo: Schaub 1985: 172; Croft 1995) there are significant correlations between 14 noun classes and semantic categories, but none of them is absolute. The meanings ascribed to some reconstructed Proto-Bantu noun classes by Denny and Creider (1986: 232–9) are shown in Table 2.1 (a somewhat different version is given in Table 11.3). In the Bantuist tradition, every countable noun is assigned to two classes: one singular and one plural (see §1.5 on terminology and §10.1.1 on the correlations between noun classes and number).

TABLE 2.1. *Semantics of noun classes in Proto-Bantu*

Noun class (SG/PL)	Semantics
Class 1/2	Human, person
Class 3/4	Extended (long) (e.g. body, river)
Class 5/6	Fruits; non-extended (e.g. stone, spot, nose)
Class 7/8	Utilitarian artefacts, despised objects and beings
Class 9/10	Animal
Class 14/6	Differentiated internal structure (e.g. bridge, bow, canoe)

In modern Bantu languages, however, noun class assignment is often much less semantically motivated, though the semantic ‘nucleus’ is still discernible. Thus, in Babungo, class 1/2 is basically human; however, it is a much bigger class than it was in Proto-Bantu, and also contains many animals, some birds and insects, body parts, plants, household and other objects, e.g. necklace, pot, book, rainbow (Schaub 1985: 175) (also see Tables 11.4 and 11.5, and Diagram 11.2).

It has often been stated that there is no real semantic basis for gender assignment of the better-known Indo-European languages. However, in a seminal study, Zubin and Köpcke (1986) provided a semantic rationale for gender assignment of nouns of different semantic groups in German (see (D) in §11.2.1). Masculine and feminine genders mark the terms for male and female adults of each species of domestic and game animals (following the ‘natural sex’ principle), and neuter is assigned to non-sex-specific generic and juvenile terms. Masculine gender is used for types of cloth, of precipitation and wind, and of minerals. Types of knowledge and discipline have feminine gender, and games and types of metal—with the exception of alloys—have neuter gender.⁵

⁵ Paul (1972) also demonstrated a partial semantic motivation for the gender assignment of English borrowings into German; for instance, drinks are mostly masculine, while fruits and flowers are feminine.

2.3.2. Morphological assignment

A connection between derivational suffix and noun class can form a morphological basis for noun class assignment. In German at least some derivational affixes are each associated with one gender, e.g. *-ung* 'action noun' is feminine and *-chen* 'diminutive' is neuter (Zubin and Köpcke 1986; Plank 1986). In Portuguese, a number of derivational suffixes (e.g. *-ção* 'action noun', as in *marca-ção* 'marking') indicate feminine gender.

In a language with a number of nominal declensions, each may correlate with a gender. In Russian the semantic assignment is restricted to human and higher animate referents. Otherwise, gender assignment is linked to declension: all nouns of declension 1 are masculine, nouns of declensions 2 and 3 are feminine, and all the rest are neuter (Corbett 1991: 40).

2.3.3. Phonological assignment

No noun class system in the world is assigned by phonological principles only. The application of phonological principles of assignment is usually restricted to nouns with inanimate referents. In some languages every noun which ends or begins with a certain vowel or consonant must belong to a particular gender. In Qafar (Saho-Afar, East Cushitic: Corbett 1991: 51–2) nouns with inanimate reference whose citation form ends in a vowel are feminine; all the rest are masculine. In Hausa all non-sex-differentiable nouns which end in *-aa* are feminine. In Katcha (Kordofanian: Heine 1982a: 200), any noun—unless it has a male referent—belongs to the feminine gender if it begins with *m-*. Phonological noun class assignment is found in Limilngan (Australian: Harvey forthcoming: §3.2): nouns whose initial segment is /l/ or /d/ tend to be assigned to Class 2 (which includes animals), and nouns with the initial /m/ tend to be assigned to Class 3 (which covers plants), even if their reference lies outside the semantic domain of these classes.

2.3.4. Mixed principles of assignment

No system of noun classes is completely devoid of semantic motivation. If a language has non-semantic principles of noun class assignment the assignment principles will be mixed, since there is always a 'core' where semantics operates. This 'core' includes humans in some languages and animates in others (see §11.2.1).

In the Harar dialect of Oromo (East Cushitic: Clamons 1993: 271) nouns referring to females are feminine, and noun referring to males are masculine. Nouns referring to inanimates, or animates for which sex is not important, are feminine if they end in a non-low vowel; otherwise they

are masculine. Russian has predominantly morphological gender assignment (§2.3.2). However, sex-differentiable nouns are assigned gender according to their semantics, and not their form. Thus, nouns like *mužčina* 'man' or *detina* 'big man' belong to the second declension, and should be feminine; but in fact they are masculine as far as agreement goes.⁶

An interesting interaction of semantic and phonological principles is found in Yimas (Sepik, Papua New Guinea, Foley 1986: 86 ff.; 1991), where the first four classes are assigned by their semantics: I—human males; II—human females; III—animals; class IV—culturally important plants. Classes 5–11 are motivated phonologically: the agreeing constituent repeats the last consonant of the nominal root.

Arapesh languages (Torricelli phylum, Papua New Guinea) appear to have a similar system. The assignment of two human classes, which comprise males and females respectively, is considered semantic by most scholars. Nekitel (1985; 1986; forthcoming) has convincingly argued against a 'purely phonological' assignment of noun classes in Abu' Arapesh (cf. Fortune 1942 and the analysis of Muhiang Arapesh by Conrad 1978: 92).⁷ A strong argument in favour of the semantic assignment of the human classes comes from loans. Most nouns which belong to Class 1 'masculine' contain either a final or an initial segment *n* (e.g. *aleman* 'man', *Nekitel* 'male name'). However, words like *Spiritu Santu* 'Holy Spirit' and *ankelo* 'angel' are attributed to Class 1 'masculine', although they do not contain the 'marker' *n*; similarly, *siste* 'nun' (from English *sister*) is assigned to the Class 2 'feminine' (though it does not contain *ʔ*- or *kw*-, initial sounds typical of this class). Morphological and phonological assignment may be hard to distinguish. There is a strong tendency to assign nouns which end in *-n* (e.g. *aun* 'moon') to the masculine class (unless they have a female referent, e.g. *nes* 'nurse', a loan from English); this is sometimes interpreted as a fossilized affix (cf. Conrad 1996).

Many languages display more complicated mixed principles of 'assignment' mingling semantic, morphological, and phonological criteria. Iraqw (South Cushitic; Heine 1982a: 200) has masculine and feminine genders. All nouns denoting singular male and female animates and male and female agentive nouns are masculine and feminine respectively. Singular

⁶ Loan words provide a few more exceptions to morphological assignment in Russian. *Kofe* 'coffee' is indeclinable, ends in *-e*, and does not denote a sex-differentiable being; these are usual properties of neuter nouns. The archaic form of this word (borrowed from Dutch or English: Vasmer 1953) was *kofej*, and it was assigned masculine gender since nouns ending in *-ej* are masculine; later it became *kofe*. In prescriptive Russian grammar, this noun is still considered masculine, while in substandard colloquial language it triggers neuter agreement.

⁷ For an attempt to account for Arapesh noun classes as an exclusively phonologically based system, see Aronoff (1991; 1994).

nouns ending in *-mo* and *-aw* are masculine. Nouns derived from Class I verbs are masculine while those from Class II verbs are feminine.⁸

A more 'grammatical' system of gender assignment can move towards a more 'semantically' oriented one. In Cantabrian Spanish a number of semantic features have been introduced for the assignment of inanimates to masculine or feminine gender (Holmquist 1991: 69), e.g. feature large/small: masc. *montón* 'stack of hay' vs fem. *montona* 'very big stack of hay'; feature deprecatory/approbratory, neutral: masc. *carreteru* 'a bad road' vs. fem. *carretera* 'a road'; feature coarse/smooth: masc. *espinu* 'mountain thorn, dark coarse bark' vs. fem. *espina* 'mountain thorn, light smooth bark'—see Table 2.2. These oppositions are mostly absent from standard Spanish.

TABLE 2.2. *Semantic features for the gender assignment of inanimate nouns in Cantabrian Spanish*

Masculine	Feminine
Male	Female
Small	Large
Narrow	Wide
Vertical	Horizontal
Tall	Squat
Phallic	Supine
Coarse	Smooth
Dark	Light
Deprecatory	Approbratory, neutral
Derived	Primary
Occasional	Familiar

For nouns with animate reference, semantic assignment often overrides morphological or other principles. In Alamblak (Lower Sepik) (Bruce 1984: 97) all nouns which denote females and short, squat, or wide objects are feminine and can have a form marked with a feminine suffix. There is one exception, the word for canoe, *doh-t*, which has a feminine marker *-t*, but is always treated as masculine in agreement, in accordance with its typical 'masculine-like' slender dimensions.⁹

⁸ A similar system is found in Punjabi (Indo-Aryan: Bantia 1993: 216–18). In Chechen and Ingush (Nakh-Daghestanian: Nichols 1989a) the distribution of nouns into five classes is partially semantically motivated; there is also a dependency between class assignment and the initial consonant of the noun. Semantic and phonological principles of noun class assignment interact in Bowili (Togo Remnant, Eastern Ghana: Heine 1982a: 199–200). See also Aikhenvald and Green (1998) for a detailed description of mixed semantic and phonological principles of gender assignment in Palikur, a North Arawak language from Brazil.

⁹ The word *Mädchen* 'girl' in German may be considered an exception to this. It is neuter since it contains a diminutive suffix *-chen* which, like other diminutive suffixes, determines the assignment of neuter. However, the situation in the modern spoken language is more

Semantic, morphological, and phonological principles account for the great majority of the assignment of nouns to gender classes in any given language, but there will often be a small residue of unexplained exceptions. Modern Hebrew has two genders, masculine and feminine. The principles of assignment are semantic and morphological. Sex-differentiable nouns are assigned gender in agreement with their semantics; nouns which contain suffixes *-t* and *-a* are feminine. Nouns which refer to cities and land, and paired and some non-paired body parts, are feminine. There are a few nouns which belong to feminine gender as unexplainable exceptions: *even* 'stone', *kos* 'goblet', *eš* 'fire' (Aikhenvald 1990: 44).¹⁰

Nouns which display a conflict between different rules of gender/noun class assignment are called HYBRIDS (Corbett 1991). In Russian *mužčina* 'man' is feminine by its morphology (it belongs to the 2nd declension: see §2.3.2) but masculine according to semantics. Semantics often takes precedence in agreement, as it does here. In Russian, in the case of most nouns denoting professions which are morphologically masculine (they belong to the first declension), such as *professor* 'professor', the agreement is feminine when focusing on the female sex of the person. In Portuguese, some nouns denoting professional occupations which end in *-a* or *-e* can also be assigned either gender depending on the sex of the referent, e.g. *dentista* 'dentist', *estudante* 'student'. The issue of variable noun class assignment and agreement is taken up in §2.4.3.

2.4. Noun classes and agreement

The presence of agreement is a definitional property of noun classes which distinguishes them from a number of other noun categorization devices. The presence of agreement is linked to the opposition between inflection and derivation. A working definition of agreement is given in §2.4.1. Principles of noun class agreement are discussed in §2.4.2. Variability in agreement, and variable noun class assignment are considered in §2.4.3. Then, in §2.4.4, I discuss the problem of how to determine the number of noun classes within a language.

2.4.1. A working definition of agreement and agreement properties

Agreement is defined by Matthews (1997: 12) as a 'syntactic relation between words and phrases which are compatible, in a given construction,

complicated: *Mädchen* may trigger feminine agreement with a relative clause marker and with a possessive pronoun; neuter is preferred when the antecedent is a child rather than a grown-up girl (Fritz Serzisko, Geoff Haig, Helma Pasch p.c.).

¹⁰ See also Clamons (1995) on semantic residue in gender assignment in Oromo, a Cushitic language.

by virtue of inflections carried by at least one of them'.¹¹ Agreement implies a systematic covariance between the grammatical meanings of grammatical morphemes (cf. Durie 1986). In a study of agreement systems the questions to be answered are:

(A) Domain of agreement

What elements agree with what elements in what grammatical configurations?

It is useful to distinguish two basic types of agreement domain: (a) agreement within an NP between modifiers and heads (head-modifier type), and (b) agreement within a clause between a predicate and its arguments (predicate-argument type).¹²

Morphosyntactic loci on which agreement markers appear are called AGREEMENT TARGETS (see Corbett 1991). Noun classes can have a noun phrase, and/or a clause, as their domain of agreement.

(B) Features and principles of agreement

In what grammatical properties do grammatical elements agree and how is it marked?

The principles of agreement are linked to the assignment of agreement features. The assignment may be either purely semantic, or mixed semantic and syntactic (cf. the distinction between 'grammatical' agreement and 'notional' agreement in Matthews 1997). These issues are discussed in §2.4.2. Conditions which may allow neutralization or variation in the agreement, and limitations on the agreement, are considered in §2.4.3.

¹¹ A similar definition of agreement is provided by Steele (1978: 610): 'The term *agreement* commonly refers to some systematic covariance between a semantic or formal property of one element and a formal property of another. For example, adjectives may take some formal indication of the number and gender of the noun they modify.' See further attempts at defining agreement by Keenan (1978: 167); Lehmann (1982: 203; 1988); further analysis of the basic parameters in terms of which agreement phenomena can or should be characterized is given by Barlow and Ferguson (1988a: 3); also see Lapointe (1985: 84), and discussion in Anderson (1992: 103–18). Agreement can be taken in a wider sense to include the so-called anaphoric agreement, i.e. the 'determination of the form of personal and relative pronouns' by their antecedents (Corbett 1991: 112; cf. Lehmann 1982). Barlow (1992) has shown that there are no reasons to make a sharp distinction between agreement within a noun phrase, and antecedent-anaphora relations. Historically, grammatical agreement often comes from grammaticalized anaphoric markers (see §13.8; also see Givón 1976; Bresnan and McChombo 1986). A number of languages, including English, distinguish different forms of personal pronouns conditioned by the gender and animacy of the antecedent. If agreement is understood in a wider sense, English can be considered a language with genders (as it was done by Corbett 1991: 112, 169).

¹² Following the distinction in Anderson (1992: 106 ff.). The important difference between Anderson's approach, and the one suggested here lies in the treatment of adpositional and possessive constructions. For the reasons which will become obvious in the course of this chapter, I will consider the principles of agreement within a noun phrase under head-modifier type; and the principles of agreement within a clause under predicate-argument type (unlike Anderson, who groups together agreement of verb with its arguments, and agreement in possessive and adpositional constructions).

Some languages have different noun class/gender agreement systems depending on the domain of agreement (head-modifier vs. predicate-argument) and on the morphological class of the agreeing element. These systems, called 'split agreement', are discussed in §2.7.

Correlations between noun classes and other grammatical categories which may also influence the ways agreement operates are considered in Chapter 10; correlations with discourse-pragmatic functions are discussed in Chapter 12.

An important distinction in the morphology of many languages is that between inflectional and derivational processes. These are summarized in Table 2.3 (cf. Payne 1990; Anderson 1992: 77 ff.; Aikhenvald forthcoming d).

TABLE 2.3. *Inflection and derivation*

Inflection	Derivation
1. Usually obligatory	Optional
2. Final process (if affix, on rim of word)	Pre-final process (if affix, between root and inflection)
3. Forms a complete word	Derives a stem which takes inflections
4. Defining characteristic of a word class (e.g. nouns inflect for case)	Usually specific to a word class
5. Do not change word class	May derive a stem of a different word class, or may add some semantic specification to a root without changing class
6. May indicate grammatical relationship between words, and/or participate in agreement	Never indicate grammatical relationship between words or participate in agreement
7. Tend to be smallish systems	May be large systems
8. Tend to have high frequency in language	Likely to have lower frequency
9. Tend to be monosyllabic likely to undergo phonological processes when combined with stem (such as assimilation, or fusion)	May be longer and are less likely to undergo phonological processes

By virtue of being realized as agreement markers, noun classes have to be treated as an inflectional category. Note that noun classes marked on the head noun ('head classes': see §2.6.1, and Evans 1997) can have a derivational function, since they may derive a stem of a different word class, e.g. Kikongo (Bantu) Ø/ba¹³-bakála (CL1/2-male) 'man', ki-bakála

¹³ Note that here and in other examples from Bantu languages the two numbers correspond to singular and plural class markers.