

The Paradigmatic Structure of Person Marking

The Paradigmatic Structure of Person Marking

Een wetenschappelijke proeve op het gebied van de Letteren

Proefschrift

ter verkrijging van de graad van doctor
aan de Katholieke Universiteit Nijmegen,
volgens besluit van het College van Decanen
in het openbaar te verdedigen op
maandag 29 januari 2001,
des namiddags om 1.30 uur precies

door

Michael Alexander Cysouw
geboren op 18 juni 1970 te Nijmegen

Promotor: Prof. Dr. P.A.M. Seuren

Co-promotor: Dr. L. Stassen

Manuscriptcommissie: Prof. Dr. H. Steinhauer (voorzitter)

Prof. Dr. E. König (Freie Universität Berlin)

Prof. Dr. G.G. Corbett (University of Surrey, Guildford)

Contents

Long Contents	VII
Acknowledgments	XIII
List of Abbreviations and Symbols	XV
1. Introduction	1
Part One: Person Marking	
2. One among the crowd	35
3. Group marking	63
Part Two: Paradigmatic Structure	
4. The diversity of the core	99
5. Paradigmatic explicitness	163
6. Connecting paradigms	197
Part Three: Number Incorporated	
7. Cardinality	223
8. The diversity of minimal number	241
9. Cognate paradigms revisited	281
10. Finale	311
Appendices	
A. Compound pronouns	325
B. Survey of person paradigms	341
C. Survey of number marking	345
References	347
Samenvatting	371
Curriculum vitae	381

Long Contents

ACKNOWLEDGMENTS	XIII
LIST OF ABBREVIATIONS AND SYMBOLS	XV
CHAPTER 1 INTRODUCTION.....	1
<i>Objective, definitions, method and some history</i>	
1.1 THE FEAT OF DOMINGO DE SANTO TOMÁS	1
1.2 DEFINITIONS AND DELIMITATION	4
1.2.1 Preamble	4
1.2.2 Person and number	5
1.2.3 Paradigmatic structure	7
1.2.4 Specialisation and grammaticalisation	10
1.2.5 Remaining delimitations	13
1.3 METHODOLOGY	17
1.3.1 Preamble	17
1.3.2 Typology versus cross-linguistic research	18
1.3.3 Sampling in typology	19
1.3.4 Crypto-diachrony	20
1.4 PREVIOUS CROSS-LINGUISTIC INVESTIGATIONS	21
1.4.1 Preamble	21
1.4.2 Forchheimer and his critics	21
1.4.3 Greenberg and his co-workers	24
1.4.4 Mühlhäusler and Harré	26
1.4.5 Laycock on New Guinean pronouns	27
1.4.6 Conclusion	28
1.5 OUTLINE OF THE BOOK	28
PART ONE PERSON MARKING	31
CHAPTER 2 ONE AMONG THE CROWD.....	35
<i>The marking of singular participants</i>	
2.1 INTRODUCTION	35
2.2 DEFINITION OF SINGULAR	36
2.3 TYPOLOGY	39
2.4 SINGULAR HOMOPHONY	40
2.4.1 Preamble	40
2.4.2 Dutch-type homophony (Sb)	41
2.4.3 Spanish-type homophony (Sc)	43
2.4.4 English-type homophony (Sd)	46
2.4.5 Summary	48

2.5	SINGULAR HOMOPHONY AND <i>PRO-DROP</i>	51
2.6	ZEROS IN PARADIGMS WITHOUT HOMOPHONY	53
2.6.1	Preamble	53
2.6.2	Zero speaker or addressee	53
2.6.3	Zero other	56
2.6.4	Non-existing other	58
2.7	CONCLUSION	60
CHAPTER 3 GROUP MARKING		63
<i>Redefining plurality in the pronominal domain</i>		
3.1	INTRODUCTION	63
3.2	DEFINITION	64
3.3	TERMINOLOGY	66
3.4	TOWARDS A TYPOLOGY OF GROUPS	70
3.5	A PARTIAL TYPOLOGY: THE FIRST PERSON COMPLEX	73
3.6	DIFFERENT KINDS OF ‘WE’	75
3.6.1	Type (Pa): ‘unified-we’	75
3.6.2	Type (Pb): ‘no-we’	77
3.6.3	Type (Pc): ‘only-inclusive’	79
3.6.4	Type (Pd): ‘inclusive/exclusive’	80
3.6.5	Type (Pe): ‘minimal/augmented’	81
3.6.6	Rare types	86
3.7	GENERALISATIONS	90
3.8	CONCLUSION	93
PART TWO PARADIGMATIC STRUCTURE		95
CHAPTER 4 THE DIVERSITY OF THE CORE.....		99
<i>A survey of possible patterns for singular and group marking</i>		
4.1	INTRODUCTION	99
4.2	METHOD OF CLASSIFICATION	101
4.3	VARIANTS OF THE ‘UNIFIED-WE’ TYPE	104
4.3.1	Preamble	104
4.3.2	Latin-type paradigm	106
4.3.3	Sinhalese-type paradigm	108
4.3.4	Berik-type paradigm	110
4.3.5	Rare variants of split ‘unified-we’	113
4.3.6	The 1/2-homophony of ‘unified-we’	116
4.3.7	The 1/3-homophony of ‘unified-we’	119
4.3.8	The 2/3-homophony of ‘unified-we’	122
4.3.9	The unified-homophony	125
4.3.10	Summary	126
4.4	VARIANTS OF THE ‘NO-WE’ TYPE	128
4.4.1	Preamble	128
4.4.2	Split variants of ‘no-we’	129
4.4.3	Homophonous variants of ‘no-we’	134
4.4.4	Summary	136

4.5	VARIANTS OF THE ‘ONLY-INCLUSIVE’ TYPE	137
4.5.1	Preamble	137
4.5.2	Sierra Popoluca-type paradigm	137
4.5.3	Rare variants of ‘only-inclusive’	141
4.5.4	Summary	143
4.6	VARIANTS OF THE ‘INCLUSIVE/EXCLUSIVE’ TYPE	144
4.6.1	Preamble	144
4.6.2	Mandara-type paradigm	144
4.6.3	Tupí Guaraní-type paradigm	147
4.6.4	Kwakiutl-type paradigm	148
4.6.5	Rare variants of ‘inclusive/exclusive’	150
4.6.6	Summary	152
4.7	VARIANTS OF THE ‘MINIMAL/AUGMENTED’ TYPE	154
4.7.1	Preamble	154
4.7.2	Maranao-type paradigm	154
4.7.3	Rare variants of ‘minimal/augmented’	156
4.7.4	Summary	157
4.8	CONCLUSION	158
CHAPTER 5 PARADIGMATIC EXPLICITNESS		163
<i>Structural analysis of oppositions in the person paradigms</i>		
5.1	INTRODUCTION	163
5.2	METHOD AND TERMINOLOGY	163
5.3	VERTICAL HOMOPHONY	166
5.3.1	Preamble	166
5.3.2	Some restrictions on vertical homophony	167
5.3.3	The internal structure of vertical homophony	170
5.3.4	Summary	172
5.4	SINGULAR HOMOPHONY	173
5.4.1	Preamble	173
5.4.2	Some restrictions on singular homophony	173
5.4.3	Summary	175
5.5	EXPLICITNESS HIERARCHY	175
5.5.1	Preamble	175
5.5.2	The building blocks of the hierarchy	176
5.5.3	The hierarchy	178
5.5.4	Summary	180
5.6	HORIZONTAL HOMOPHONY	181
5.6.1	Preamble	181
5.6.2	A hierarchy of horizontal homophony	181
5.6.3	Horizontal homophony and the Explicitness Hierarchy	183
5.6.4	Summary	184
5.7	BUILDING A SEMANTIC MAP	185
5.8	CORRELATIONS WITH MORPHOLOGICAL STATUS	188
5.8.1	Preamble	188
5.8.2	Explicitness Hierarchy	189
5.8.3	Horizontal Homophony Hierarchy	191
5.8.4	Amount of morphemes	192
5.8.5	Explanations	193
5.9	CONCLUSION	194

CHAPTER 6 CONNECTING PARADIGMS 197*Person paradigms through time and space*

6.1	INTRODUCTION	197
6.2	COGNATE PARADIGMS	198
6.3	UP AND DOWN THE HORIZONTAL HOMOPHONY HIERARCHY	201
	6.3.1 Preamble	201
	6.3.2 Interconnecting the small paradigms	202
	6.3.3 Up the hierarchy	204
	6.3.4 Summary	208
6.4	UP AND DOWN THE EXPLICITNESS HIERARCHY	209
	6.4.1 Preamble	209
	6.4.2 Minimal/augmented inclusive (or not)	210
	6.4.3 Inclusive/exclusive opposition (or not)	213
	6.4.4 Vertical and singular homophony (or not)	214
	6.4.5 Summary	216
6.5	CONCLUSION	216

PART THREE NUMBER INCORPORATED 219**CHAPTER 7 CARDINALITY 223***Redefining number in the pronominal domain*

7.1	INTRODUCTION	223
7.2	A META-LANGUAGE FOR DUAL MARKING	225
7.3	MARKEDNESS REVERSALS	229
7.4	OTHER NUMBERS	233
7.5	CONCLUSION	238

CHAPTER 8 THE DIVERSITY OF MINIMAL NUMBER 241*A typology of dual marking in the pronominal domain*

8.1	INTRODUCTION	241
8.2	METHOD AND TERMINOLOGY	242
8.3	TWO DIFFERENT FORMS FOR 'WE'	243
	8.3.1 Preamble	243
	8.3.2 The dual-unified-we paradigm	243
	8.3.3 'Incomplete' dual-unified-we paradigms	245
	8.3.4 Summary	249
8.4	THREE DIFFERENT FORMS FOR 'WE'	250
	8.4.1 Preamble	250
	8.4.2 Dual in exclusive only	251
	8.4.3 Dual in inclusive only	253
	8.4.4 Inclusive-exclusive in dual only	253
	8.4.5 Inclusive-exclusive in plural only	254
	8.4.6 The case of Burarra	256
	8.4.7 Summary	256
8.5	FOUR DIFFERENT FORMS FOR 'WE'	257
	8.5.1 Preamble	257

8.5.2	The dual-inclusive/exclusive paradigm	257
8.5.3	'Incomplete' dual-inclusive/exclusive paradigms	259
8.5.4	The partial-unit-augmented paradigm	262
8.5.5	The case of Palaung	263
8.5.6	Summary	264
8.6	FIVE DIFFERENT FORMS FOR 'WE'	264
8.7	OTHER DUALS	268
8.8	STRUCTURAL ANALYSIS	270
8.8.1	Preamble	270
8.8.2	Vertical homophony and the Explicitness Hierarchy	271
8.8.3	Horizontal homophony	273
8.8.4	Summary	275
8.9	CONCLUSION	276
CHAPTER 9 COGNATE PARADIGMS REVISITED		281
<i>Connecting the dual through time and space</i>		
9.1	INTRODUCTION	281
9.2	METHOD OF COMPARISON	281
9.3	LINKING THE MAJOR DUAL PARADIGMS	284
9.3.1	Preamble	284
9.3.2	Cognate paradigms	285
9.3.3	Summary	291
9.4	MINIMAL/AUGMENTED AND ITS VARIANTS	291
9.4.1	Preamble	291
9.4.2	Cognate paradigms	292
9.4.3	Summary	296
9.5	DUAL-3WE AS AN INTERMEDIATE	297
9.5.1	Preamble	297
9.5.2	Cognate paradigms	297
9.5.3	Summary	302
9.6	NUMBER MARKING INCORPORATED	303
9.7	HIERARCHIES	305
9.8	CONCLUSION	308
CHAPTER 10 FINALE		311
<i>Summary and prospects</i>		
10.1	SUMMARY	311
10.1.1	Preliminaries	311
10.1.2	Person and number	312
10.1.3	Paradigmatic structure	313
10.1.4	Hierarchies	314
10.1.5	The concept of Pure Person	316
10.1.6	Diachronic interpretation	318
10.2	PROSPECTS	319
10.2.1	Where to go from here	319
10.2.2	Gender in person paradigms	320
10.2.3	Asymmetry of affixation	321
10.3	WIDER APPLICATION OF RESULTS	323

APPENDIX A COMPOUND PRONOUNS.....	325
<i>Other person categories disqualified</i>	
A.1 INTRODUCTION	325
A.2 FROM COMPOUND TO PRONOUN	327
A.3 THE BANTOID COMPOUND PRONOUNS	330
A.4 GENERALISATIONS	337
A.5 THE INCORPORATIVE PLURAL REVISITED	338
A.6 CONCLUSION	340
APPENDIX B SURVEY OF PERSON PARADIGMS	341
APPENDIX C SURVEY OF NUMBER MARKING.....	345
REFERENCES	347
SAMENVATTING	371
CURRICULUM VITAE.....	381

Acknowledgments

This thesis would not have been written without the help of many people. Above all, my parents did not only make it possible in an existential sense, they also surrounded me with an environment of intellectual openness. I will never forget the bulky birthday present, which turned out to be Hofstadter's *Gödel, Escher, Bach*. This book has been a major stimulus for a scientific career.

At university, the inspiring classes by Pieter Seuren persuaded me to switch from my initial love Mathematics to reading Linguistics. Subsequently, I became hooked on the world's linguistic variation through the fascinating ideas of Leon Stassen.

Daily life at the institute for General Linguistics and Dialectology has been greatly enriched by the various meetings with fellow AiO's and other colleagues, be it in the context of ProZA, POEZie, at coffeebreaks (with biscuits!), lunch or dinner. In particular, I want to thank Dick Smakman, Marie-José Palmén and Lisanne Teunissen for proofreading.

As for the wider scientific community, I thank everybody at the various LOT schools for making them so much fun, especially Dirk Janssen and Christine Erb (who will both accompany me to the former East Germany) and a big 'thank you' to all colleagues around the world that have answered my questions about some far-away language.

In the meantime, life went on outside university. I have had a wonderful home with wonderful roommates at the Annastraat over the years, thank you all very much for sharing these years with me. And without all you tangodancers, acrobats, jugglers, game fanatics, puzzle designers and other friends-in-action, I would have probably lost my mind in the word processor! Yet, the greatest joy has been to share my life, both inside and outside university, with Paula-Irene Villa. Thank you for your love, help and inspiration.

List of Abbreviations and Symbols

1	speaker
2	addressee
3	other
/	or
,	and
-	morpheme boundary
+	combinations of persons into groups
→	agent/patient combinations (also used for horizontal homophony)
ASSOC	associative
DEB	debitative
CLIT	clitic
COM	comitative
COND	conditional
COP	copula
DEM	demonstrative
DL	dual
EMPH	emphatic
EXCL	exclusive (= 1+3)
FEM	feminine
FUT	future
IMP	imperative
IMPERF	imperfective
INCL	inclusive (= 1+2 and 1+2+3 combined)
IRR	irrealis
LOC	locative
LOG	logophoric
MASC	masculine
NAME	personal name
PAST	past
PERF	perfective
PLACE	geographic name
PLUR	plural
POSS	possessive
PRES	present
PRON	independent pronoun
PURP	purposive
REPET	repetative
SG	singular
TAM	not further specified tense/aspect/mood marker
WH	question word

Chapter 1

Introduction

Objective, definitions, method and some history

Haiku, you ku, he
She, or it kus, we ku, you
Ku, they ku. Thang ku.

Ted Hipple, *The traditional grammarian as poet*
(cited in Wales, 1996:1)

1.1 The feat of Domingo de Santo Tomás

The history of linguistic investigation is full of discoveries. One of those discoveries was made by a Spanish friar known by the name of Domingo de Santo Tomás. By way of introduction, the story of his insight will be told here to illustrate the theme and the approach of the present investigation.¹

In 1540, the Dominican friar Domingo de Santo Tomás set sail from Spain on his way to Perú. He belonged to a group of missionaries that were recruited by Francisco Martínéz Toscano to work in this newly discovered part of South America. At that moment, only seven years after the defeat of the Inca empire by Francisco Pizarro, the journey to this new Spanish territory was full of perils and uncertainty. Yet, Domingo de Santo Tomás survived the trip and managed to do his missionary work among the native people of Perú. His catechistic work made him an expert in the ‘lengua general’, the general language spoken in the region, nowadays referred to as Quechua. In

¹ The (sparse) biographical materials on Domingo de Santo Tomás are from Rodolfo Cerron-Palomino, as added to the 1994 facsimile edition of the original text of de Santo Tomás (1560).

1555, after 15 years of missionary work, Domingo de Santo Tomás returned to Spain, determined to do a load of work. Among many other things, he had taken up the task to inform the authorities in Spain about the miserable situation of the native people of Perú. Further, he wanted to take care in person of the printing of his grammar-cum-dictionary of the Quechua language. This resulted in the *Grammática o arte de la lengua general de los Índios de los Reynos del Peru*, which appeared in 1560, printed in Valladolid in an astonishingly large edition of 1500 copies. The outline of this Quechua grammar follows the traditional descriptions of the grammar of Latin. However, there are numerous details where the Latin format does not suffice for the intricacies of the Quechua language. In many aspects, this language is different from Latin and the classical descriptive format had to be enhanced to explain the structure of Quechua. One of the inventions that Domingo de Santo Tomás introduces in his grammar is an explanation of the difference between the two forms for ‘we’ that are found in Quechua. He explains that one of the two forms is used for ‘we, including the person or persons that are spoken to’, and the other form is used for ‘we, excluding the person or persons with whom we are talking’. This difference, grammaticalised in Quechua, can be made explicit in English by adding extra linguistic material, like *we all here present* for the inclusive meaning, or *we, me and my friends* for the exclusive meaning. Today, almost 450 years after the publication of this explanation, the analysis is still considered accurate. Even the terminology has remained; this opposition is known today as a difference between an ‘inclusive’ and an ‘exclusive’ first person plural.

‘El plural deste pronombre *ñóca* es *ñocánchic* o *ñocáyco*, que quiere dezir ‘nosotros’. Y es de notar que entre *ñocánchic* y *ñocáyco* ay dos diferencias, una intrínseca, de parte de la significación dellos, otra extrínseca, de parte del verbo que les corresponde ... La primera es que, aunque *ñocánchic* y *ñocáyco* significan ‘nosotros’, el *ñocánchic* significa ‘nosotros’, connotando & incluyendo en sí la persona con quien hablamos: como hablando con indios, si quisiésemos dar a entender que también entran ellos y se incluyen en la habla que hablamos con este pronombre *nosotros*, como diciendo: a nosotros nos crió Dios, usaríamos de este pronombre *ñocánchic*, esto es ‘nosotros, incluyendo también los indios’. Pero si los quisiésemos excluir a ellos de la tal razón o plática ... en este lengua (por ser más abundante) no ay necesidad de añadir determinación alguna sino solamente usar de este pronombre *ñocáyco*, que quiere dezir ‘nosotros’, connotando que se excluye de aquella pluralidad la persona o personas con quien hablamos.’ (de Santo Tomás, 1560:8-9)

[The plural of this pronoun *ñóca* is *ñocánchic* or *ñocáyco*, which mean ‘we’. It is to be noted that between *ñocánchic* and *ñocáyco* there are two differences, one intrinsic, due to their meaning, the other extrinsic, due to the verb that corresponds to them. ... The first is that although *ñocánchic* and *ñocáyco* both mean ‘we’, *ñocánchic* means ‘we’, with the connotation of including the person as such with whom we are speaking; like when speaking with Indians, when we want to give to understand that they also take part, and that they are included in the meaning of what we say with this pronoun *we*; as when one would say: ‘God created us’, we will use that pronoun, *ñocánchic*, that is, ‘we’, including also the Indians. But if we want to exclude them from the meaning or the speech ... in this language (to be more explicit) it is not necessary to add any clarification, except to use the pronoun *ñocáyco*, meaning ‘we’, with the connotation of excluding the person or persons with whom we are talking from the plurality.]

The merits of this discovery are not to be underestimated. Domingo de Santo Tomás had to work out this aspect of the grammatical system of Quechua from scratch. The

grammatical description of languages other than Latin or Greek was still in its infancy.² The difference between an inclusive and an exclusive variant of the word ‘we’ was unheard of, even inconceivable, in the occidental tradition of grammatical analysis. None of the major languages in Europe and in the Near and Middle East show such an opposition. The nearest languages (from a European perspective) which have an inclusive/exclusive difference are some tucked-away Caucasian languages, the Dravidian languages all the way down in southern India, and some Nilo-Saharan languages in the sub-Saharan savannah of Africa. As it turned out, the Quechua-speaking people in far away South America were the first to direct Western attention to this peculiar possibility of human language. Yet, once the difference between the two forms of ‘we’ is explained, it appears to be a completely natural distinction. After some time, one even starts wondering how it is possible to do without this very practical device in linguistic interaction.³

The story of de Santo Tomás illustrates the basic impetus for cross-linguistic research, which is the method that will be followed in the present work. Cross-linguistic research attempts to learn about the possibilities of human language by studying the various ways in which communities around the world put their linguistic competence to work. A comparison of the structure of very diverse languages presents an opportunity to escape the entrapment of the (linguistic) imagination. The scientific lore about the extent of variation of human language is strongly biased by the structure of presently known languages. Certain possibilities of human language would never be conceived if it were not for the fact that they happen to exist. It was the existence of an overtly marked inclusive/exclusive opposition in Quechua that opened the eyes of the Western linguistic tradition to the possibility of such categories. In contrast, other possibilities might be thought of as logically possible, but might turn out to be unattested. In this study, I will use the variety of the world’s linguistic structures to sketch an outline of the possibilities of human language – within a restricted domain.⁴ The story of Domingo de Santo Tomás also illustrates the domain of the present study. As can be learned from Quechua, some languages use two referentially different words for the group of people that includes the speaker (ie the English word ‘we’). In this study, I want to address the question of what possibilities human languages use to mark the participants in a speech-act. Thus formulated, this is a rather extensive domain of inquiry. In all languages, there are many ways to refer to ‘me, myself and I’. One finds expressions like ‘the undersigned’ or ‘the present author’; or simply ‘Michael’ or ‘Mr. Cysouw’; not to speak of the numerous other epithets that could be

² The first printed grammar other than of Latin or Greek is the Spanish grammar by Lebrija from 1492. In total, Rowe (1974) counted only 10 languages other than Latin and Greek of which printed grammars appeared before the Quechua grammar by Domingo de Santo Tomás –who was probably unaware of these other descriptions, except for the Spanish grammar of Lebrija.

³ More details on the history of the discovery of the inclusive/exclusive opposition can be found in Haas (1969), Hardman (1972), Mannheim (1982), Suárez Roca (1992) and Adelaar (1993).

⁴ If a particular logically possible structure is not attested among this world’s languages, it is not necessary to categorise it as ‘impossible’. The actual languages of this world do not necessarily exhaust the possibilities of linguistic structure. However, the more languages are taken into the comparison, the unlikelier it becomes for a non-attested structure to be an actual possibility for human language. Still, it might be better to talk about ‘improbable’ instead of ‘impossible’ when talking about non-attested structures in the world’s linguistic diversity.

used to refer to this specific person. Once the full expressive power of any natural language is taken into account, the possibilities to mark participants in a speech-act become innumerable. Given the right context, almost every kind of word in a language can be used to refer to a participant. To remain within workable limits, this investigation will only deal with an analysis of the PARADIGMATIC STRUCTURE of SPECIALISED person marking. This definition of the subject introduces two important restrictions on the domain of investigation. First, the person markers have to be grammaticalised as a specialised part of a language. The markers have to be specifically devoted to person marking. The variety of referential possibilities is restricted in this study to those linguistic elements that do nothing else besides person marking, as, for example, the English words ‘I’ and ‘you’. Second, such specialised person markers are investigated from the structural context of the paradigm. Specialised person markers are normally found to be part of a closed set of person markers that are treated identically within the structure of a language. Such a set of equivalent markers is called a ‘paradigm’. The precise referential value of a person marker can only be specified in relation to the other person markers in the paradigm.

In this introductory chapter, miscellaneous definitional, methodological and historical aspects are discussed. I suggest that the reader who is mainly interested in the actual cross-linguistic comparison skips the considerations in this introductory chapter. At the beginning of each chapter, the specific definitions, methodological issues and results of earlier research are discussed extensively. In this chapter, addressed to the more-than-casual reader, a survey of the approach that is taken in the present study is given. First, the definitions and delimitation of the domain are made explicit in section 1.2. Subsequently, some methodological issues will be dealt with in section 1.3. Next, in section 1.4, previous cross-linguistic work on person marking is discussed. Finally, section 1.5 offers a survey of the structure of the book and leads to the core of the present work: the cross-linguistic variation of the paradigmatic structure of person marking.

1.2 Definitions and delimitation

1.2.1 Preamble

Language is a communicational device. Reference to the participants in a communicational setting can be marked linguistically in various ways. A father who is talking to his baby might refer to himself by using the word ‘daddy’ as in *daddy is busy now*. In English, such self-reference by using a full noun is a marked construction, mainly to be heard in motherese. However, in other languages, like Thai, this usage is much more widespread, to be found throughout the various registers of the language (Cooke, 1968:44-55). In all languages, many different words can be used for self-reference, although in some languages it will take more conversational preparation to make the self-referring usage of a word intelligible. The problem is that words with self-reference change their referential value depending on the person who is speaking. For example, the understanding of such shifting reference is known to present major difficulties to the language-learning child. Jespersen (1922:123) tells a delightful story of two playing children, one of whom does not get the shifting reference of the

word ‘enemy’. We can envision the incredulous child shouting: ‘*I am the enemy, not you!*’ Jespersen calls such linguistic elements ‘shifters’. All linguistic elements that are used in a deictic (‘pointing’) function are shifters. Locational deixis (*here, there*), time deixis (*now, tomorrow*) and participant deixis (*me, you*) all involve shifting reference of linguistic elements (Jespersen, 1922: 123-124; Bühler, 1934: 79-148; Jakobson, 1971; Hengeveld, 1997).

Amidst all these shifters, this study will deal only with specialised shifters used for reference to speech-act participants. To be included in the investigation, the shifters have to be ‘specialised’. This means that the shifters do not have any other possible usage besides being a shifter. The words *here, now* and *me* are such specialised shifters. In contrast, words like *daddy* or *enemy* can be used as shifters, but do not necessarily function as such. Second, this investigation will deal only with shifters that are used for reference to participants in the speech act. Basically, the speech-act dyad of speaker and addressee will be the starting point for the present investigation. Summarising, there are three criteria for linguistic elements to be included in the investigation: they have to be a shifter, specialised for that function, and used for reference to speech-act participants. Henceforth, linguistic elements that adhere to these three criteria are called PERSON MARKERS (see further section 1.2.2).

Person markers appear almost universally within the confines of a paradigm. Person markers, as defined above, do not stand alone within a language. They are part of a closed set of person markers that have an identical place in the structure of a language. Such a set of elements is called a PARADIGM. This investigation addresses the question to what extent the structure of person marking paradigms is found to vary cross-linguistically. The survey of the variability of the paradigm structure can be read as a prolegomenon to a theory of the ‘richness’ of a paradigm (see further section 1.2.3). Diachronically, person markers do not behave differently from other linguistic elements. They grammaticalise from independent nouns into person markers. Also, independent pronouns grammaticalise into inflectional person markers. All specialised participant shifters will be included in this study, irrespective of independent or inflectional morphology. Moreover, all person marking ‘agreement’ is included. The inflation of the word ‘agreement’ and the often assumed primacy of morphologically independent person markers is countered (see section 1.2.4). Finally, some borderline cases and some problems with the cross-linguistic delimitation of the domain will be discussed (see section 1.2.5).

1.2.2 Person and number

‘Thus the speaker, the listener, and the things spoken about are three essential factors of normal speech. To these must now be added the actual words themselves.’ (Gardiner, 1932:28)

This study is concerned with specialised shifters that are used for reference to speech-act participants. The (extra-linguistic) setting of a speech act invokes a few salient cognitive categories. The principal categories of participant deixis are ‘speaker’ (the originator of the speech) and the ‘addressee’ (the recipient of the speech). All deixis that does not include either of these categories can be summarised negatively. The ‘other’ participant is everything (not necessarily human or animate) that is neither

speaker nor addressee. These principal categories will form the basis of the present investigation. The specialised linguistic elements that code these categories are called person markers. Following the conventions of the occidental grammatical tradition, the coding for speaker will be called ‘first person’; the coding for addressee will be called ‘second person’; and the coding for any other participant will be called ‘third person’. The linguistic coding of these three categories will be discussed extensively in chapter 2.

There is more to a conversation than only speaker and addressee. In an analysis of conversational practices, Goffman (1979) argues for a decomposition of the notions ‘speaker’ and ‘addressee’. The traditional notion ‘speaker’, he pleads, is a cluster of various communicational functions, like ‘animator’ (ie ‘the sounding box’), ‘author’ (ie ‘the agent who scripts the lines’) and ‘principal’ (ie ‘the party to whose position the words attest’) (Goffman 1979:16-17, cf Levinson 1988:169). Likewise, the traditional notion of ‘addressee’ is to be decomposed into, at least, the functions ‘hearer’, ‘unaddressed’, ‘over-hearer’, ‘bystander’ and ‘eavesdropper’ (Goffman 1979:8-9). In a thorough reappraisal of Goffman’s proposals, Levinson (1988:170-180) further systematises the decomposition of ‘speaker’ and ‘addressee’. There are various reasons why one would want to have a more finely grained description of the possible functions of a participant. The most promising area of application is the ethnography of speaking, as Levinson shows (1988:192-221). However, for the analysis of grammatical structure, the merits of the decomposition of the traditional speech-act roles are less obvious. The basic opposition between speaker and addressee is rather commonly grammaticalised among the world’s languages. In contrast, none of the finer grained role distinctions are attested as grammaticalised categories.

‘The classical analysis has held up remarkably well in the face of recent comparative analysis. The great majority of languages exhibit the three persons in a paradigm of pronouns, verb agreement, or elsewhere.’ (Levinson, 1988:182-183)

The predominant grammaticalisation of the traditional referential categories ‘speaker’ and ‘addressee’ is probably a result of functional pressures, Levinson argues. Prototypically, language is an interactional device with two speech-act participants who organise their speech in a turn-taking sequence. The speaker/addressee roles in the turn-taking system are prototypically reflected in the structure of language.

‘Quite probably, the universal tendency in languages to distinguish, in pronominal categories or elsewhere, primarily and prototypically the two deictic categories of first and second person, is related closely to the superordinate categories of speaker and addressee/recipient that are the basis of the turn-taking system.’ (Levinson, 1988:176,183)

However, by proposing this explanation, Levinson counters Goffman’s original motive for a finer grained distribution of referential categories. Of course, Goffman argues, the prototypical speech-act situation is a dyadic one, but many other configurations are attested, and are probably just as common in daily speech-interaction.

‘The ratified hearer in two-person talk is necessarily also the *addressed* one, that is, the one to whom the speaker addresses his visual attention and to whom, incidentally, he expects to turn over the speaking role. But obviously two-person encounters, however common, are not the only kind; three or more official participants are often found. In such cases it will often be feasible for the current speaker to address his remarks to the circle as a whole, encompassing all his hearers in his glance, according them something like equal status. But, more likely, the speaker will, at least during periods of his talk,

address his remarks to one listener, so that among official hearers one must distinguish the addressed recipient form *unaddressed* ones.’ (Goffman, 1979:9)

Goffman argues convincingly that much more is possible, but eventually, it is an empirical question whether these possibilities are also used in the grammatical structure of language. As it appears now, this is not the case. Earlier research did not come up with any example, and during my investigations for the present work, I have also not come across any new claims in this direction. Languages do not use more finely grained interactional categories in their grammatical structure. Still, it is important to keep open the possibility that grammatical markers have simply not been recognised as such. For the present investigation, I will disregard any of these more finely grained categories and restrict myself to the traditional gross categories ‘speaker’ and ‘addressee’.

On the basis of the principal categories ‘speaker’ and ‘addressee’, groups of participants can be formed. Groups of participants consist of more than one participant and are thus necessarily plural. For example, speaker and addressee can make up a group, or speaker and other, or speaker and addressee and other, et cetera. All theoretical possible combinations will be discussed extensively in chapter 3. In that chapter, it will also be argued that the notion ‘plural’ is not appropriate for these groups of participants. Groups of participants consist of more than one individual, so they are plural by definition. However, they are plural in a completely different sense from that in which normal nouns can be plural. As will be shown later in chapter 7, the notion of ‘number marking’ within the domain of person marking is better reserved for categories that are traditionally known as ‘dual’, ‘trial’, et cetera. Finally, there is a recurrent claim in the literature that there are some languages with a special set of person markers, variously called ‘compound pronouns’ or ‘complex pronouns’. This phenomenon will be discussed in appendix A. It will be shown there why I have chosen to disregard these pronouns in the main line of the present work.

1.2.3 Paradigmatic structure

Grammaticalised person markers do not occur in isolation within a language. They preferably appear in a PARADIGM of person markers. The structure of a person marking paradigm is the central aspect of the present inquiry. I will investigate the structure of paradigms that show at least an opposition between speaker and addressee in the marking of singular participants. Some elucidation of the notion of a paradigm in the present work will be given in this section.

First, a central assumption is that a person marking paradigm is a closed class of linguistic elements that occur in complementary distribution. Stated differently, a paradigm is a set of linguistic elements that occur in the same syntagmatic place in the structure of a language. The syntagmatic-paradigmatic duality is taken in direct reference to the way it was introduced by de Saussure (1916:170-180).⁵ For example, the subject pronouns of English constitute a paradigm, as they occur in complementary

⁵ In the literature, the syntagmatic/paradigmatic duality is rightfully connected to de Saussure. However, de Saussure uses the term ‘associatif’ instead of the expected term ‘paradigmatique’. He reserved the word ‘paradigme’ for the traditional notion of declension (cf de Saussure, 1916:15).

distribution as subject of a predicate. The object form of the pronouns belong to a different paradigm because these forms are used in a different syntagmatic place in the language.

Second, there is no impossible reference in a paradigm. That is to say, the mutual exclusive elements in a person marking paradigm fill out the complete referential array of possible participants. For example, singular reference is often divided into ‘speaker’, ‘addressee’ and a third class. This third class consists of all reference that is not either of the first two. This third ‘non-person’ simply fills all possibilities that are not taken care of by the other elements in the paradigm. A similar tendency is attested for plural marking. If there is no specialised element for any plural reference, then the missing referential value is taken over by one of the other elements in the paradigm. Preferably, a person marking paradigm is ‘referentially complete’.⁶ It is not the language in which ‘tout se tient’, but the paradigm.

Third, I will interpret a paradigm as a closed space of alternative options. The individual person markers in a paradigm do not arrive at their referential value intrinsically, but in mutual delimitation to the other elements in the paradigm. The space of reference to all possible participants is subdivided into referential categories by the available elements in the paradigm. The referential space that is allotted to a specific element in a paradigm can only be defined in relation to the other elements. This perspective of the paradigm is reminiscent of the division of vowel space by the available phonemic vowels. A low open vowel /a/ is a rather different vowel when it is found in a three vowel system as compared to when it is found in a seven vowel system. Another analogous concept to the paradigm is the ‘Wortfeld’ in lexicography, as proposed by Trier (1931), in which the meaning of a word is dependent on the meaning of other closely related words.

Fourth, the present study will be concerned with the internal structure of person-paradigms. Such a paradigm is only a small part of a complete language. Normally, there are multiple person marking paradigms within a single language. English has independent pronouns – both in a nominative and in an oblique form; then there are possessive pronouns and finally there is a paradigm of person markers in the present inflection. All in all, there are at least four different paradigms that share the referential work. I have not attempted to formulate a criterion that would set apart one of the various paradigms as the primary one. Not the languages as wholes, but the individual paradigms within each language are the crux of the comparison. The cross-linguistic comparison in this investigation is on a sub-language level (cf the ‘item-based’ approach from Nettle, 1999). Consequently, a particular language can appear multiple times in this investigation; each time with different paradigms. The result of this approach is an insight into the paradigmatic structure of person marking. Only indirectly, will this help to understand the functioning of a whole language.⁷

⁶ The word ‘preferably’ has been added because there are some exceptions, as there are always exceptions to a cross-linguistic generalisation. Specifically, problems arise in the case of a paradigm with ‘zero’ independent pronouns (see section 2.6). In a few extreme cases, there are only two independent pronouns (one for speaker and one for addressee), and all other reference is ‘zero’ (see section 4.4). Probably, it is better to consider these ‘zeros’ as non-existing. Such paradigms with non-existing marking are not referentially complete.

⁷ The importance of a proper study into the paradigmatic structure can, for example, be inferred from the generative analysis of *pro*-drop. In this proposal, the concept of ‘richness’ of a paradigmatic

Combining this perspective of paradigmatic structure with the perspective of person marking as explained in the previous section, I have arrived at the following definition of the object of my investigation. I will investigate the structure of paradigms that show at least an opposition between speaker and addressee in the marking of singular participants. There are a few borderline issues concerning this definition. First, there are cases that do not show an opposition between speaker and addressee, but that have an opposition between speaker/addressee and other. These cases are included in the investigation. Second, there are cases that have an opposition between speaker and addressee, but that have an irregular paradigmatic structure. These cases are excluded from this investigation.

The first of these issues ('speaker/addressee' versus 'other' opposition) is important for the classification of the English inflection. The English present inflection is included in this study as a borderline case of a person marking paradigm. Following the definition of person marking as outlined in the previous section, the English inflection should not be included because there is no opposition between the form of (*I*) *speak* versus (*you*) *speak*. In both cases the inflection is zero. However, the English inflection has an opposition between speaker/addressee and any other singular participant: (*I/you*) *speak* versus (*he/she/it*) *speaks*. I have decided to interpret this opposition as a borderline case of person marking. In contrast, the (spoken) French inflection does not have any opposition in the singular: (*je*) *parl(e)* versus (*tu*) *parl(es)* versus (*il/elle/on*) *parl(e)*. In all three forms, the inflection is zero. Speaker and addressee are not marked in the French inflection, not even in combined opposition to any other participant. The decision to include structures like the English inflection, but exclude structures like the French inflection is a rather ad-hoc limitation that arose because I have taken the marking of singular marking as the basis of the classification (see chapter 2).⁸

The second issue is important for the classification of a few troublesome cases in which the marking of the speaker and addressee is syntagmatically different, yet both markers belong to a single paradigm. The worst troublemakers are paradigms in which the first person is marked by a suffix and the second person is marked by a prefix, or vice versa. An example of such a paradigm is found in the Berber languages. The subject inflection in Berber is a suffix in the first person singular, but a prefix in the first person plural. Second person marking is a combination of prefixes and suffixes.⁹ Another instance of a combination of prefixes and suffixes is attested in the Muskogean languages of North America. The first person agent marking in the

[continued from previous page]

structure is crucial (Harbert, 1995:222). However, I have not been able to find a precise definition in the literature that explicates under what conditions a paradigm is to be considered 'rich'. The present investigation, specifically chapter 4 and chapter 5, can be read as an attempt to formulate a cross-linguistically sensible definition of the 'richness' of a person marking paradigm. The resulting notion of 'explicitness' of a person marking paradigm does not coincide with the presumed notion of 'richness' from the generative literature (see also section 2.5).

⁸ Except for French, I know of only a few incidental cases that have been excluded from this investigation because of this decision (see footnote 7 on page 40).

⁹ See, for example, the Berber language Tamazight (Willms, 1972:196). This structure of Berber is reminiscent of (and probably diachronically related to) the imperfect affixes of the Semitic languages, as, for example, described for Cairene colloquial Arabic by Gary (1982:100). The Arabic imperfect is discontinuous. Most of the reference is marked by prefixes, but gender and part of the number reference are disambiguated by suffixes.

Muskogean languages is a suffix, but all other person markers are prefixes.¹⁰ Finally, the object marking in the Tequistlatecan languages of Mexico shows different affixation for first and second person. A first person object is marked as a prefix while all other persons are marked as suffixes.¹¹ To conclude, paradigms with a mix of prefixes and suffixes for speaker and addressee reference occur, although they are not common from a world-wide perspective. The few examples just mentioned are disregarded in the present investigation. Less severe problems are the paradigms in which speaker and addressee reference is syntagmatically identical (say, both are prefixes) but some other material in the paradigm is syntagmatically different (say, the first person plural is a suffix). In such cases, I will restrict the extent of the paradigmatic structure to those elements that are syntagmatically identical to the speaker and addressee (in the hypothetical example, this means that I only include prefixes into the paradigmatic structure). An example will illustrate this restriction of the paradigmatic structure. An inflected verb from Upper Bal (a dialect of the South Caucasian language Svan) is shown in (1.1). The paradigm presented is the imperfect of a verb of Class A (Tuite, 1997:28). Aspects of person marking are attested both as prefixes and as suffixes.

(1.1) UPPER BAL

<i>1 Sg:</i>	xw-ama:r-äs	<i>1 Incl:</i>	l-ama:r-ad
		<i>1 Excl:</i>	xw-ama:r-ad
<i>2 Sg:</i>	x-ama:r-äs	<i>2 Plur:</i>	x-ama:r-ad
<i>3 Sg:</i>	∅-ama:r-a	<i>3 Plur:</i>	∅-ama:r-a

The person marking from Upper Bal is analysed as two different paradigms, one prefixal and one suffixal. The prefixal paradigm is the most genuine person marking device, distinguishing speaker from addressee (and even separating an inclusive). The suffixal paradigm is a borderline case of person marking. There is no opposition between speaker and addressee reference, but speaker and addressee are marked by a special morpheme different from other, and different from plural participants (cf the English inflection).

(1.2) UPPER BAL

<i>Prefixal paradigm:</i>		<i>Suffixal paradigm:</i>	
		<i>Singular</i>	<i>Plural</i>
<i>1 Inclusive:</i>	l-...		
<i>1 Exclusive:</i>	xw-...		
<i>2:</i>	x-...	...-äs	...-ad
<i>3:</i>	∅-...	...-a	

1.2.4 Specialisation and grammaticalisation

Only person markers that are specialised shifters are considered in this study. Independent pronouns like *I* or *you* in English are such specialised shifters: the reference

¹⁰ This structure is found throughout the Muskogean family. For data on Chickasaw, see Payne (1982:359); for data on Alabama, see Lupardus (1982:66-74); for data on Koasati, see Kimball (1985:107) and for data on Choctaw, see Nicklas (1974:31). Haas (1977) proposes an origin of the person marking by grammaticalisation of auxiliaries.

¹¹ For data on Highland Chontal, see Turner (1966:65,68) and for data on Huameltultec Chontal, see Waterhouse (1967:356).

of these words has to shift with the speaking participant. It is impossible for them not to shift.¹² In contrast, words like *daddy* can be used as a shifter, but are not necessarily used in that sense. Such words are excluded from the investigation. For languages like English, this exclusion still leaves the specialised shifters. However, in some languages, it appears to be the case that nothing remains when all non-specialised shifters have been removed. This situation seems to occur in many Southeast Asian languages. For example, Cooke (1968) has made an extensive investigation of person marking in Thai, Burmese and Vietnamese. The amount of possibilities is overwhelming. The usage of proper names, kin terms and normal nouns for speaker or addressee reference is completely normal.¹³ For Thai, Cooke mentions 27 more or less specialised first person markers and 22 second person markers (1968: 11-18). Moreover, these are only a tiny portion of the possibilities as ‘not even personal pronouns are listed in their entirety in each language, much less other pronominally used forms’ (Cooke, 1968:2). And then it also turns out that most of the ‘specialised’ person markers are still interpretable in the original (non-shifting) meaning. The shifters mean, for example, something like ‘individual’, ‘crown of the head’ or ‘servant’. The least one can say, is that ‘real’ person markers are a mystifying category in Thai. I would even say that ‘real’ person markers do not exist. Of course, it is perfectly possible to perform self-reference in Thai; it is even possible to do so in many more ways than in English. However, there is not one predominant way to do so. In order not to confuse the analysis more than is necessary, I have excluded the person marking from languages like Thai from the present investigation.¹⁴ Probably, specialisation of shifters can occur by grammaticalisation from erstwhile nouns, as found in Thai.¹⁵ However, it turns out to be rather difficult to find good examples of this grammaticalisation. Once the person markers are specialised, the original nominal meaning is long gone. A fairly good case is the Classical Malay noun *sahaya*, meaning both ‘servant’ and ‘I-humble/polite’, which turned into modern Indonesian *saya*, meaning ‘I-neutral/non familiar’. The word *sahaya* no longer exists in Indonesian and *saya* does not mean ‘servant’. In this case the semi-grammaticalised shifter *sahaya* turned into the completely grammaticalised shifter *saya* (H. Steinhauer, personal communication).

Another, well documented, aspect of the grammaticalisation of person markers is from independent pronouns to inflectional person marking (often called ‘agreement’).¹⁶ Givón (1976) argues explicitly for this continuum to be interpreted as

¹² The only possibility for the English pronoun *I* not to shift with the speaker is in a meta-linguistic use, as when talking about a novel one says: ‘*It is the ‘I’ of the novel that is the murderer!*’

¹³ Somewhat less exotic, this was also found in Swedish, ‘where until comparatively recently title and name were used in place, or in avoidance of pronominal address’ (Howe 1996: 11, citing Mårtensson 1988: 143).

¹⁴ For a comparable system, though somewhat less elaborate than Thai, see Japanese (Shibatani, 1990:371-388). Often, if a language in Southeast Asia has a specialised set of person markers, these ‘real’ person markers are only very sparingly used, as, for example, in Mandarin Chinese (Li & Thompson, 1979).

¹⁵ Cf von Humboldt (1830) and Blake (1934).

¹⁶ There is much literature on the notion of agreement which tries to cope with the large variation of agreement-like features among the world’s languages (Moravcsik, 1978; Corbett, 1979; Lehmann, 1982; 1983; Moravcsik, 1988; Corbett, 1994). The provisional conclusion is that agreement is about

showing that there is no structural difference between the two forms of person marking. Person can be marked independently or inflectionally. The reasons for a particular language to choose either of these possibilities remains unclear. However, from a cross-linguistic perspective, both strategies are not *a priori* different.

‘... the tacit assumption [is] that agreement and pronominalization are two distinct processes. I will suggest below that they are fundamentally one and the same phenomenon, and that neither diachronically nor, most often, synchronically could one draw a demarcating line on any principled grounds.’ (Givón, 1976:151)

Still, in most of the generative literature on pronominal marking, this continuum is not accepted. Independent pronouns and inflectional person marking are considered to be two completely different aspects of linguistic marking. The reason for the persistence is probably the high status of the ‘projection principle’ as formulated by Chomsky (1981:29). This principle proposes a strong constraint on syntactic analyses and the intermediate transformations. It implies that at every level of syntactic analysis, the arguments of each predicate are to be present (overt or covert). Independent pronouns are possible instantiations of arguments; inflectional person marking is seen as agreement of the predicate with these arguments. Thus, a regular noun or an independent pronoun has to be present at each level of syntactic analysis. There has been an attempt to lessen the power of this principle by allowing argument status to inflectional person marking (or clitics), and consider independent pronouns as optional adjuncts. This idea was originally proposed by Jelinek (1984).

‘I argue that the clitic pronouns do not constitute agreement (AGR) with a nominal ... My claim will be that verbal argument arrays (argument positions) in LS are satisfied always and only in PS in Warlpiri by clitic pronouns, and that nominals are simply optional adjuncts, with non-argumental functions.’ (Jelinek, 1984:44)

Many a word has been written on the argumental status of inflectional person markers since then. For example, Baker (1990; 1991) attempted to improve on the original proposals. Others, like Saxon (1986:142) in a description of Dogrib, oppose these proposals. From a more descriptive point of view, some fuel has been thrown into the discussion by Mithun (1986; 1991). Finally, the debate seems to have ended after the publication of an article by Austin and Bresnan (1996) which shows that Jelinek’s analysis of Warlpiri does not hold for close relatives of Warlpiri.

‘The clitic pronouns that Jelinek (1984) and others take to be the source of non-configurationality in Warlpiri are simply an areal feature of Australian languages that is independent of the characteristics of free word order, null anaphora, and split NP’s. They do not provide the unifying explanation for non-configurationality of the Warlpiri type.’ (Austin & Bresnan, 1996:263)

Notwithstanding all the commotion, the idea that inflectional person markers could be more important for the structure of a language than its independent pronouns does not seem to have found its way into the generative canon. Generative textbooks like (Webelhuth, 1995) and (Culicover, 1997) do not mention any of the arguments, nor cite any of the literature dealing with this proposal. The independent pronouns are still

[continued from previous page]

covariation between two linguistic elements. Whether one of the two is controlling the other differs from construction to construction, and from language to language.

‘the term *agreement* commonly refers to some systematic covariation between a semantic or formal property of one element and a formal property of another.’ (Corbett, 1994, citing Steele 1978)

considered to be the principal person marking system and the term ‘agreement’ remains to be used practically as a synonym of inflectional person marking in most of the syntactic literature.¹⁷

In contrast, I will follow Givón’s proposal, and treat independent pronouns and inflectional person marking as two different, though *a priori* equivalent ways of person marking. Whether there is a difference between the two strategies is not clear beforehand. I will include both inflectional person markers and independent pronouns in my search for the variability of paradigmatic structure. I include (in principle) each and every person marking paradigm of a language in the investigation without restrictions as to the function of the paradigm within the language.¹⁸ I will not choose one paradigm per language as representing the ‘real’ person marking system of the language.

1.2.5 Remaining delimitations

The objective of this book includes the marking of person in the singular and the non-singular. The restriction to the combination person-number may sound natural to many readers, but from a cross-linguistic perspective the separation of these two dimensions from other dimensions is far from obvious. With the risk of raising unnecessary suspicion, I want to show why it is not instantly clear how this domain can be delimited by presenting a few examples of pronominal paradigms that cross the boundary of person and number. The following selection of examples and references is intended to show the inherent variability of human language. Even this seemingly clearly distinct domain of linguistic structure will turn out to have no sharp edges.¹⁹

The first and most far-ranging restriction that is made is that I will disregard person marking paradigms that distinguish gender. I will not only disregard the gendered elements in the paradigms and retain the rest, but I will disregard the whole paradigm when it contains gendered forms. The reason for this rather crude decision is that I do not know which form to leave out when there are various gender forms. In fact, a separate chapter on the interrelation between gender and person marking was planned, but did not materialise. For more information on the interrelation between gender and

¹⁷ Interestingly, the proposal that inflection is for some languages better considered as the ‘principal’ person marking system had already been developed by the classical Arabian grammarians from the 8th century AD onward (Versteegh, 1997:77-81). They ran into trouble in the 10th century AD when they tried to reconcile Aristotelian logic (which was based on a separable predicate-argument structure) with their traditional grammatical analysis (Abed, 1991).

¹⁸ In fact, there is one restriction on the paradigms included. Only person marking paradigms are considered that are used as argument in a clause or as one-word phrases. The other major grammatical construction with person markers, viz pronominal possession, is disregarded in this study. The reason for this restriction is a practical one. The original impetus for compiling the cross-linguistic data was to look for the differences in argument status of person markers. It was only after most of the data-collection and much of the analysis had been done that I decided that is impossible to analyse the argument status of a person marker without first analysing the structure of the paradigm. To control whether the omission of pronominal possession influences the conclusions of this investigation, I have had a look at the structure of such paradigms. This look has been more than just a spurious glance, although it has not been representative enough to be reported on here. My impression is that the omission does not influence the conclusions as put forward in this work. The only place where a different conclusion might be reached, concerns zero marking for speaker-reference (see section 2.6.2).

¹⁹ For a comparable survey of possible categories in person systems, see Mühhäusler & Harré (1990:60-88).

person, see Corbett (1991:126-132) and Plank & Schellinger (1997). Only in a few exceptional cases, when an omission would result in the disappearance of particular rare paradigmatic structure from the investigation, have I retained paradigms that maximally distinguish one or two gender forms in the third person singular. The eventual occurrence of gender marking is always explicitly noted in the discussion of the examples. The complete variation of gender marking in the domain of person marking deserves more than just a spurious glance. To raise the interest for this subject, I will illustrate some of the more curious interrelations between gender and person with two examples. The first example is from Barasano, an eastern Tucanoan language spoken in Peru. The gender/person suffixes in Barasano mark gender and animacy in the third person. Curiously, the inanimate suffix is also used for speech-act participants.²⁰

‘Agreement of person, number, gender and animacy between subject and verb is required, and is shown by subject agreement markers which occur in final position in the phrase. ... *-ha* is used for all inanimates, and for speech-act participants, i.e. first or second person, singular or plural.’ (Jones & Jones, 1991:73-74)

The second unusual structure is found in the Afro-Asiatic stock. It is widespread throughout the Semitic, Cushitic and Berber languages. In these languages, the pronominal prefixes often have the same form for the second person and for the third person feminine. The morpheme for the third person masculine is different. In the Cushitic languages, this third person masculine is often identical with the first person singular; an ‘interlocking’ pattern, as it is called by Tucker (1966: 15-16).

Another major omission of the present work is that I will disregard honorific usage of pronouns and specialised politeness forms of pronouns (Head, 1978; Mühlhäusler & Harré, 1990:131-167; Simon, 1999). This is a tricky delimitation, as it might be difficult to distinguish the ‘real’ referential usage from the ‘metaphorical’ polite usage of a pronoun. As it turns out, most descriptions are quite clear on the distinction between referential value and honorific usage. Also from conversational analysis, it appears that the referential and the metaphorical instances of pronoun use are distinguishable (Schegloff, 1996:441-449). I will only present a single – somewhat unusual – case of honorific usage of pronouns to exemplify the intricate complications that can arise in this dimension. This interesting ‘metaphorical’ usage of person distinctions is found in Galela, a West Papuan language spoken on Halmahera in Indonesia. In this language, the difference between inclusive and exclusive first person plural is mixed with honorific usage. In specific contexts, the inclusive form is used as a polite variant of the first person exclusive.

‘... a situation as where a man asks members of another family to give money to his own family ... His own family is referred to by *na-* [inclusive] which is interpreted to mean exclusive, polite. Sociolinguistically, what happens is that by including the addressee grammatically, the man relieves his embarrassment for having to ask for money.’ (Shelden, 1991:165-166)

A different kind of pronominal ‘politeness’ is attested in Australia. Some Australian languages have specialised person markers that have to be used when referring to kin.

²⁰ See for a comparable structure, though in a completely different part of the world, the usage of zero gender marking in Archi, a Nakh-Dagestanian language from the Caucasus (Corbett, 1991:127-128).

For example, in the Mamu-dialect of Dyirbal, a specialised pronoun is attested to refer to the combination ‘self and spouse’ (Dixon, 1972:50-51). Specialised kin-related person markers are most prolifically found in the Arandic family, for example in the languages Alywerre (Koch, 1997) and Mparntwe Arrernte (Wilkins, 1989: 126-128). In these languages, complete sets of specialised pronouns exist, in which reference depends on patrimoiety and generation.

Other traditional *pro*-nominal elements are disregarded, like interrogative pronouns, indefinite pronouns, reflexive pronouns and reciprocal pronouns. However, in many languages, the reflexive and reciprocal markers belong structurally to the person marking paradigm. For example, the object paradigms in the Athabascan languages include specialised reflexive and reciprocal morphemes.²¹ Another example of the fluid division between reflexive, reciprocal and person marking is found in Mundani, a Grassfields language from Cameroon. In Mundani, the object pronoun is used for person reference, but also for reflexive and reciprocal marking. Structurally, there is no difference between the three meanings of the Mundani sentence as presented below.

(1.3) MUNDANI

<i>bóó'</i>	<i>n'tiŋá</i>	<i>áwóó</i>
3PL,PRON,SUBJ	help	3PL,PRON,OBJ
‘They are helping them’		
‘They are helping themselves’		
‘They are helping each other’		

(Parker, 1986: 149)

In some person paradigms, the spatial configuration of the referents is specified in more detail. An example of spatial specification is attested in the paradigm of independent pronouns from Cuna, a Chibchan language from Panama. All independent pronouns are characterised by a suffix *...-i*. However, this suffix is also added to the demonstrative roots, distinguishing four grades of locational specificity: ‘this’, ‘the same’, ‘that’ and ‘yonder’. These demonstrative roots are morpho-phonetically included into the person marking paradigm (Holmer, 1946: 190). Another variant of spatial specification in the person marking paradigm is attested in the Uto-Aztec language Ute. In the third person, an opposition between visible and non-visible is grammaticalised (Givón, 1984: 356-357). Finally, a combination of spatial specification and marking of visibility is found in Kwakiutl (Boas, 1947: 252). A variant of the spatial specification is a discourse specification. Instead of marking the place of the object in ‘real’ space, the place of the referent in the space of the discourse can be specified by specialised pronominal forms. A specific way of grammaticalising discourse specification is by so-called ‘obviative’ third person markers. Obviative markers indicate that the intended referent is not the participant which is most prominent in the discourse; it is ‘the other one’ that is intended. Obviative marking is most famous as a characteristic of the Athabascan languages like Slave (Rice, 1989: 431) and Navaho (Young & Morgan, 1987: 8). It is, however, widespread throughout North

²¹ For data on Slave, see Rice (1989:431) and for data on Navaho, see Young (1987:8).

American languages, as in the Algonquian languages, for example in Menomini (Bloomfield, 1962:44). Spatial specification in the third person is disregarded.

Some languages distinguish special logophoric pronouns. The precise extent of the usage of logophoric pronouns differs from language to language, but the general idea behind these pronouns is that they are used in quoting other people. I will disregard such pronouns. For example, Babungo, a Grassfields language spoken in Cameroon, has such specialised logophoric pronouns. The pronoun *yì* is a logophoric third person pronoun that is generally used as a subject in embedded sentences (Schaub, 1985: 111-113). The complex logophoric pronoun *vìyyì* is used for ‘semi-direct speech’. Semi-direct speech is direct speech of a participant in a story (Schaub, 1985:2). This is exemplified in example (1.4). Instead of the logophoric pronoun, also the inclusive basic pronoun *nsôo* or the third person plural pronoun *vǎŋ* could have been used (Schaub, 1985:3).

(1.4) BABUNGO

fəkáy *gí* *lāa* *síi* *kà* *vìyyì* *ỳ* *kə́*
 tortoise say that doubt can LOG do what

Tortoise said, “What can **we** do?”

(Schaub, 1985:2)

Still another aspect of person marking that will be disregarded in the present investigation is the specialised portmanteau forms for transitive constructions. In many languages around the world, sentences like ‘I see you’ are not formulated by using two separate pronouns, but by using one person marker that is specialised for this combination of subject and object. For example, in Diegueño, a Yuman language spoken in the USA, such specialised person prefixes are attested for various transitive combinations. The prefixes *?-...*, *m-...* and *ø-...* are used for the first, second and third person intransitive, respectively. Two other prefixes, *nʹ-...* and *?nʹm-...*, belong to the same paradigm, but these prefixes are used for the transitive combinations of 1subject/2object and 1object/2subject, respectively (Langdon, 1970:139-140).²² The occurrence of transitive portmanteau forms in Australian languages is extensively discussed by Heath (1984; 1991). More recently, Heath (1998) gives a quick but thorough survey of such forms in American languages. Transitive portmanteau prefixes will be disregarded in this study.

Finally, a solitary case of yet another specialised person marker that will be disregarded is found in , a Papuan language from New Guinea (Irwin, 1974:14-15). For this language, a special person suffix for ‘speaker alone’ is described. There is a regular first person suffix *...-m(in)*. This suffix is ambiguous in its reference because it is used both for singular and plural reference. In some cases, another suffix *...-l* is used to stress the fact that only the speaker alone is intended.²³

²² The actual interaction of the various prefixes and the various transitive combination of participants is more complicated than this, but this quick review of the basic situation sketches the kind of structure that is attested in the Yuman languages.

²³ Irwin (1974:14) uses the word ‘exclusive’ for this specialised ‘speaker alone’ form, but it is clear from the description that this pronoun is rather different from what is normally called an ‘exclusive’ pronoun.

This list of the many facets of person marking among the world's languages is probably far from complete. The more languages are studied, the more strange and unexpected phenomena (from a Eurocentric perspective) will crop up. Every delimitation that is proposed for a cross-linguistic study is bound to encounter exceptions and problematic cases when confronted with the actual linguistic variation. Still, in this variation it is not chaos that rules. In the remainder of this work, I will show that there are still clear tendencies and near-universal generalisations to be made. However, the balance between generalisations and exceptions is a precarious one. Forgetting the exceptions might present a picture of the structure of human language that is oversimplistic. In contrast, disregarding the gross generalisations amidst the variations might suggest a massive variability of linguistic structure that is not representative of the actual situation. I will seek a balance between the extremes.

1.3 Methodology

1.3.1 Preamble

The present investigation consists methodologically of three different kind of studies. Some parts use a cross-linguistic method, some parts use a typological method and some parts use a crypto-diachronic method. In section 1.3.2, I will discuss the difference between the cross-linguistic method and the typological method. Basically, the cross-linguistic method will be used to arrive at a firm basis for the typology. Next, in section 1.3.3, I will discuss a few special methodological aspects of the construction of the typology, specifically the question of sampling. Finally, in section 1.3.4, I will present a somewhat hybrid method, a combination of cross-linguistic and diachronic ideas, which I call 'crypto-diachronic'. This method will be used to tap into the diachronic dynamics of the paradigmatic structure of person marking. The result of this crypto-diachronic study will be a cognitive map, which links paradigmatic structures that are conceptually (and probably also diachronically) close to each other.

Before I turn to these themes, I want to speak a few words of gratitude and caution. The words of gratitude are directed to the authors of the grammatical descriptions that provide the necessary data for my investigation. The morphophonemic analysis of a language may be taken as given. The methods to identify phonemes and morphemes seems to be fairly accurate and stable. That is to say, the identification that there exists a certain morpheme in a language can normally be made with relative certainty. Also, the acknowledgment of a paradigm in a language is usually rather straightforward. But the question of the function or meaning of the morphemes is a completely different problem. I see the most important task of descriptive linguistics as a search for the best fitting and most compelling analysis of the meaning of specific morphemes (or combinations of morphemes). The present investigation builds on the large body of work compiled by the authors of such invaluable descriptions of specific languages. I am writing this book standing on the proverbial shoulders of giants. The words of caution concern the information on individual languages quoted in this work. I have attempted to quote only original sources, tracking down all references to their roots. Only in a few cases have I been unable to get hold of the original source.

In those cases, I have added the name of the author of the original source to the reference in which I found the data. The data on all various languages have been copied as accurately as possible from the original source. However, I cannot guarantee that I have in all cases represented the data in the intended format. I suggest that any reader who wants to quote data from this work, consults the original sources to prevent any accumulation of misquotation.

1.3.2 Typology versus cross-linguistic research

There is a methodological difference between constructing a typology and doing cross-linguistic research. Basically, the construction of a typology is only part of a cross-linguistic investigation. However, the non-typological side of cross-linguistic research is often taken for granted in methodological statements (pace Croft, 1990: 1-18). For the present work, the cross-linguistic investigation has been of major importance for the subsequent typology. In fact, most of the research time has been spent on cross-linguistic work, while the typology was compiled relatively quickly once the parameters had been fixed.

A TYPOLOGY is always constructed on the basis of a ‘tertium comparationis.’ A typology is senseless if it is not constrained by an explicitly formulated subject to be compared in the world’s linguistic variation. On the basis of such a tertium comparationis – which has to be formulated extra-linguistically – different languages can be categorised on how they formulate that part of linguistic structure. However, a tertium comparationis is never theory-independent. It presupposes something to be universally expressible in language. This universal hypothesis originates in the theoretical background on the basis of which the typology is made. In this strict interpretation, typology is a hypothesis-testing methodology. Given a view about the structure of human language, the world’s linguistic variation can be used to test this view by making a typology. In the present study, the hypotheses to be tested will come from a CROSS-LINGUISTIC investigation. I see cross-linguistic research as the attempt to understand language by combining different viewpoints from different languages and different grammarians into a coherent picture.²⁴ The differences that are found in grammars are not necessarily caused by the language, they could just as well be caused by the viewpoint of the grammarian.²⁵ However, as a matter of principle, I assume that all descriptions are good reflections of the language. Every grammar is probably incomplete; certain aspects are omitted, other aspects are exaggerated. However, in most cases the bias of a description is caused by a ‘bias’ in the structure of the language. A

²⁴ The philosophical background to this interpretation of cross-linguistic research is the insight that the most difficult problem for scientific inquiry is to bypass the ‘spectacles of tradition.’ People cannot observe objectively; there is always some prior experience or knowledge that interferes with the observation, colouring it and making every observation always an interpretation. One way to bypass this obstacle is to combine different viewpoints. Different viewpoint are often caused by different ‘spectacles.’ On the presupposition that the alternative perspectives look in more or less the same direction, but from different points of view or with different magnitudes of resolution, the different results obtained from alternative perspectives can be used to understand the influences of the ‘spectacles’ on the resulting picture. By combining the different views, a new level of understanding can hopefully be reached.

²⁵ In some cases, I have been able to correct clear errors in the description by comparing the data of different sources of the same or of closely related languages. Such errors are checked with the original author (if possible) and are always explicitly noted in footnotes to the data-presentation.

specific language may structurally stress certain functions of human communication, prompting a grammarian to spend more time dealing in greater depth with this particular function. The differences between the descriptions of the world's languages are better seen as a possible opening to new insights than as a nuisance that complicates comparison.

Cross-linguistic research and typology follow each other in an interpretational circle. For the cross-linguistic part of the research, traditional linguistic notions, like 'pronoun', 'first person' or 'plural', are taken as a starting point. These traditional notions, and many others, are part of an ancient heritage of meta-language, of talking about language, that have proven to be successful for that purpose. But being successful in the past does not mean that these terms transcend theory. These words are (part of) a theory about language, just as any other linguistic theory. These 'traditional' terms propose a kind of format with which one can tackle the interpretation of linguistic structures. They are words that help to understand the structure of a language, but it should be realised that they do not necessarily have to be appropriate for all languages.²⁶ In some cases, I will propose a slight reformulation of the interpretation of traditional concepts to be able to deal better with the attested cross-linguistic variation. On the basis of these revised concepts, a typology of the world's linguistic variation will be made. This typology 'tests' the revised concepts, and the results of this test can be used as a starting point for a next round of cross-linguistic research. In the present work, the chapters 2, 3 and 7 employ a cross-linguistic method. Building on this basis, the chapters 4, 5 and 8 describe the cross-linguistic diversity and present a typology.

1.3.3 Sampling in typology

The typological side of this investigation needs some more methodological explication. The typological work consists of two different parts. The first part is the compilation of the typology itself, and the second a quantitative analysis of this typology. Let me first explain how the typology itself has been compiled. In chapters 4 and 8, the various paradigmatic structures of person marking are classified into different types. These types are based on the structure of the 'first person complex'. The first person complex is the set of categories that include at least the speaker (see chapter 3). This compilation of different paradigmatic types was mainly driven by the urge to present the wide range of possibilities that is attested regarding person marking in the world's languages. The sample of languages for this compilation was not fixed beforehand. Of course, every part of the linguistic globe has been included in this typology, but this diversity-constraint has only been taken as a base-line condition, not as a guarantee for empirical success. Every single example of rare paradigmatic structures has been included in the sample to show the inherent variability of human language. The result is a broad typology that includes many cases that are only attested in one or

²⁶ With 'traditional' terminology, I refer to the same kind of theoretical assumptions as Dixon (1997:128-135) refers to by 'Basic Linguistic Theory':

'Over the past few hundred years work has been done on language from every part of the world, with many aspect of linguist theory being rethought, reformulated and refined as a result. The term Basic Linguistic Theory has recently come into use for the fundamental theoretical concepts that underlie all work in language description and change, and the postulation of general properties of human language.' (Dixon, 1997:128)

two examples. On the other hand, not every case of the commonly occurring paradigmatic structures is included. After a representative group of a specific common paradigmatic structure had been described, I stopped the collection of more exemplars of such common paradigms. This first part of the typological investigation uses a qualitative method. Two qualities are searched for: to represent the word's linguistic diversity and to show roughly which structures in this diversity are common and which are rare. On the basis of this qualitative typology, some quantitative analyses are performed in chapter 5 and in the last part of chapter 8. These analyses use the set of examples that had been compiled for the qualitative typology to perform some quantitative calculations. As a result of this decision, the sample for these analyses is not statistically constrained, as, for example, proposed by Rijkhoff & Bakker (1998). Still, I believe that the results of these quantitative analyses are valuable. The number of included cases is high, and the diversity of the sample is large. Even more strongly, the sample covers each and every structure that I know of, including all quaint cases. A methodology that is more strongly constrained by statistical representativeness would probably find less diversity, not more. This means that any correlation that holds for the present diversity-sample will surely also hold for a statistically constrained sample.

1.3.4 Crypto-diachrony

The chapters 6 and 9 have not been mentioned until now in this methodological explanation. These two chapters use yet another method to reach yet another goal. I will call the method 'crypto-diachronic', as it is a combination of cross-linguistic and diachronic ideas. The main objective of this method is to get an inkling of the similarity space of the paradigmatic structure of person marking.

The typological chapters 4, 5 and 8 have resulted in a set of paradigmatic structures that are regularly encountered among the world's languages. These paradigmatic structures are not invariable instances of linguistic form, but perpetually changing configurations of morphemes. Structures that are similar will change more easily from one to the other. From the typological investigation, a hypothesis can be distilled on the similarity of paradigmatic structure. I will use the crypto-diachronic method to test the hypothetical similarity space of paradigmatic structure. In chapters 6 and 9, I will look at groups of closely related languages that have slightly different paradigmatic structures of person markers.²⁷ The morphemes in these paradigms are individually cognate. However, the paradigmatic configuration of the morphemes is slightly different. I assume that these paradigmatic differences are caused by a relatively recent change in one direction or the other. In most cases, I do not propose a direction of change. In this sense, I am not proposing any diachronic developments. However, I conclude from such examples of recent change that the paradigmatic structures involved are similar. By compiling a large set of such cognate paradigms, I will be able to build a similarity map of paradigmatic structures. This method will only be used to investigate the hypothetical similarity space that resulted from the typological investigation.

²⁷ Cf the method of intragenetic comparison, as discussed by Greenberg (1969:184-194).

1.4 Previous cross-linguistic investigations

1.4.1 Preamble

There exists a massive literature on person marking, but the majority of these investigations deals with the analysis of the pronominal elements of a particular language or of a group of closely related languages. These studies mainly provide continuously improved analyses of the pronominal systems of these languages. In some studies, a theoretical hypothesis is confronted with the phenomena as found in the particular language or language group. In others, idiosyncratic peculiarities of the language in question bring about new grammatical concepts to describe linguistic phenomena. Yet, in contrast to this plethora of individual studies, the amount of literature that takes a cross-linguistic perspective on person marking is rather limited. In such cross-linguistic work, the variety of structures that languages use for person marking is compared in an attempt to arrive at a better understanding of what is meant with the concept ‘person marking’. I will here discuss the few cross-linguistic investigations of person marking that were published in the second half of the twentieth century. I do not discuss specific claims about cross-linguistic variation that are made by the various authors. The history of typological claims will be discussed wherever appropriate throughout the present work. In this section, only the general approach and the methodological issues of the studies are discussed.

The first extensive cross-linguistic collection and comparison of person marking was compiled by Forchheimer (1953). I will discuss his work together with some of the devastating critiques that were published at the time of publication in section 1.4.2. Some cross-linguistic claims about person marking are found in the typological work of Greenberg and his co-workers. I will review these works in section 1.4.3. Mühlhäusler and Harré (1990) wrote a monograph predominantly dealing with the deictic aspects of person marking, using data from a wide variety of languages. I will discuss this work in section 1.4.4. Finally, I will discuss a shorter work by Laycock (1977) in section 1.4.5; this investigation takes a structurally restricted perspective on person marking as Laycock only looks at independent pronouns.²⁸

1.4.2 Forchheimer and his critics

Over the last fifty years, the main reference point for anybody who wanted to know anything about pronominal systems has been Forchheimer’s thesis *The Category of Person in Language* from 1953. The enduring reference to this work stands in strong contrast to the hostile reactions uttered shortly after the book was published. Fred Householder’s review in *Language* (1955) was highly critical. He objects to the classification ‘of language phenomena in terms of some pre-established pattern of Latin

²⁸ In the generative literature, and particular in the wake of the ‘government & binding’ theory, there has been a great deal of attention for the anaphoric use of pronominal elements. I do not feel competent to review this extensive literature, which, as a whole, could be interpreted as being ‘cross-linguistic’ because many different languages are discussed by many different authors. However, the main body of work in this tradition is based on analyses of Germanic and Romance languages. There are a few attempts to reconcile the European-based analyses with phenomena found in other languages. To a relative outsider like me, the results do not look promising for a true cross-linguistic theory (cf Mühlhäusler & Harré, 1990:55-57; Huang, 1995).

grammar or some triumph of logic' (Householder, 1955:94), and he argues for a more explicit way of selecting data.²⁹ Dell Hymes' review in the *International Journal of American Linguistics* (1955) is equally trenchant. He objects to the inaccurate presentation of the data and the implicit diachronic argumentation.

'Perhaps so weak a book should not have been written; on the other hand, it is the only one in its field.' (Hymes, 1955:300)

There are indeed many points in Forchheimer's study that call for criticism. However, one can hardly expect such a ground-breaking kind of investigation to be perfect. The work by Forchheimer is a first attempt at a cross-linguistic comparison of the structure of pronominal marking. Among the many problems, I believe there to be many stimulating aspects in Forchheimer's investigation, although the correct conclusions are only covertly stated in the work itself. The main error, I believe, is the title of the book, which puts the reader (and Forchheimer himself) on a wrong track.

Forchheimer's study is truly cross-linguistic in nature. He explains that he has 'examined a good five hundred grammars and word lists. ... general works and collections on almost all areas' (1953:2), although only about 70 different languages are actually discussed in the book (cf Ingram, 1978:218). This wide array of linguistic diversity that is included in the comparison is typical of the German tradition of linguistic comparison and classification as it was found at the start of the twentieth century (cf Müller, 1876-1887; Schmidt, 1926; Royen, 1929). Unfortunately, Forchheimer followed this tradition at a time when nobody was really interested in such work. The restriction to roughly 70 languages stems from Forchheimer's primary interest in the diversity of linguistic structure. He does not list all languages that he has investigated, but only those languages that exemplify a particular type of structure.

'Wherever I have found or suspected unusual or new material I have changed from the extensive to an intensive study of that linguistic stock or area. ... after a certain stage had been reached, examination of more languages did no longer show new types.' (Forchheimer, 1953:2-3)

The types that he was interested in are various 'person patterns' or 'person systems' (Forchheimer, 1953:2). However, his main interest lies not in the paradigmatic structure as I have defined it (see section 1.2.3), but in the morphological patterns of the marking of number in the pronominal domain. This central occupation with the morphological structure can be inferred from the kind of types that are distinguished in the typology. His classification consists of types like 'Languages with Morphological Plural of Pronouns' or 'Languages without Morphological but with Lexical Plural at least in the First Person', et cetera. The result is a typology of the morphological coding of the connection between person and number (Forchheimer, 1953:11-19). This goal of Forchheimer's investigation can be directly connected to the influence of

²⁹ Householder starts off referring to Forchheimer as an amateur, indicating that he does not intend to take the work seriously.

'The linguistic amateur, like the amateur in other fields, has characteristic virtues and vices. He is usually enthusiastic, sanguine, and industrious, often learned. His goals and ideals, on the other hand, may appear strange, even futile, to his more conventional colleagues.' (Householder, 1955:93)

the work of Wilhelm Schmidt (especially 1926:316-334), an influence that is explicitly mentioned by Forchheimer in the acknowledgments to his book.

A central problem with Forchheimer's study (criticised by both Householder, 1955:94, and Hymes, 1955:295) is that he does not explicate on what grounds he selects the person-marking data from the various descriptions. He takes the term 'person marking' to be a self-explanatory term, describing a certain universally applicable concept of linguistic structure. Forchheimer implicitly took the rough equivalents of West European pronouns as the data for his study. By definition, these elements will mark for person, but the unanswered question is whether those element necessarily form a special category. He does not give any criteria on the basis of which he decides that a particular set of person marking morphemes form a category. It is surely true that every language can mark person in some way, but whether every language has a category of person remains unproven. This point leads to another problem with the implicit assumptions in Forchheimer's study. Not only does he assume a category of person in all languages, he also claims to be able to extract *the* system of person-marking for all languages:

'From any fairly reliable grammar, be it conservative or modern, prescriptive or descriptive, scholarly or written by an interested traveller, it is possible to work out *the* system of person peculiar to a language.' (Forchheimer, 1953:1, italics added)

Many languages have an independent pronoun system alongside inflectional pronominal systems. It remains unclear how Forchheimer decides which one is *the* system. Sometimes, he takes an independent person marking paradigm as *the* system. In other cases, he includes an inflectional paradigm in his sample as *the* pronominal system.

'Personal pronominal elements are found in different forms. They may occur either as enclitics (affixes), or as so-called free, separate, independent, disjunctive, or absolute pronouns. ... After studying hundreds of systems, I became convinced that the affix-pronouns often represent the pure pronominal elements.' (Forchheimer, 1953:23)

It seems that Forchheimer finds *the* system of person marking in all languages, not as an empirical result, but because of an *a priori* assumption that there has to be such a thing in all languages. In conclusion, it can be said that Forchheimer's reasoning can only be understood if the existence of a universal and unitary person system is taken for granted and not as the result of his investigation. From this discussion it should not be concluded that such an assumption is wrong altogether. It is only on the basis of assumptions that questions can be asked and results can be reported. Problems arise when the assumptions are confused with the research questions and the conclusions. This occurs, for example, when Forchheimer claims universal status for his assumed categories, as he occasionally does.

'The distinction of speaker, addressed, and neither speaker nor addressed is universally found.' (Forchheimer, 1953:39)

This conclusion is not correct, given the approach that Forchheimer uses. He *assumes* that the person categories exist; he does not *investigate* whether they actually do.³⁰

³⁰ Cf section 1.2.4 of this work and Householder (1955:96) on a negative answer to the question of whether specialised person markers exist universally.

One can regret the fact that Forchheimer took this perspective. He could indeed have performed a different investigation, but, ultimately, he did what he did.

‘... he intends to concern himself mainly with the morphological and etymological relationships of the actual forms in each language. What a lost opportunity! When he could have actually studied the grammatical category of person, found out which languages do and which do not have such a category, determined the various systems in the different languages and compare them ...’ (Householder, 1955:94)

In general, besides rather overt flaws in the representation of the data (cf Householder, 1955:97-99; Hymes, 1955:294-295) and some confusions between assumptions and results, the work is an interesting effort to do cross-linguistic research at a time when linguistics was not very interested in such work. Maybe the major mistake Forchheimer made was the title of his study: *The Category of Person in Language*. He does not investigate this theme, but assumes *a priori* that the category of person exists. It would have been better if he had called the book something like ‘*Number marking in Person Systems*’.

1.4.3 Greenberg and his co-workers

The general reception of Joseph Greenberg’s famous 1963 article on universals of grammar contrasts sharply with the hostile reaction to Forchheimer’s study. In only ten years, the scientific climate had changed drastically, as some of the same points of criticism that were raised against Forchheimer’s study can also be raised against Greenberg’s article. The most important point of criticism to be made is that Greenberg – just as Forchheimer – does not explicate on which grounds he selects his data.³¹

In the 1963 article, besides the well known universals on word order, there is also a section on morphology where Greenberg makes some claims about pronominal elements. Greenberg claims universal status for three persons and two numbers:

‘Universal 42: All languages have pronominal categories involving at least three persons and two numbers.’ (Greenberg, 1963:96)

The almost identical claim was made by Forchheimer, but there is a slight difference in phrasing between Greenberg and Forchheimer. Greenberg does not claim that there is something like *the* pronominal system. He instead argues that there are three persons and a number opposition somewhere in each language, but he does not claim a single primary paradigm with these characteristics for each language. Yet, the same criticism that was raised in the discussion of Forchheimer applies here. If the selection criterion is to find elements comparable to the European pronominal elements in other languages, one will probably always find them. The remaining problem is whether those elements form a genuine category, as seen from the point of view of the structure of the language in question.

Following Greenberg’s approach to language universals, the *Project on Language Universals* took place in Stanford between 1967 and 1976, directed by Charles

³¹ Greenberg even explicitly states that ‘no attempt at definition of categories will be attempted’ (Greenberg, 1963:92). He ignores this aspect because he wants to concentrate on other questions:

‘To have concentrated on this task, important in itself, would have, because of its arduousness, prevented me from going forward to those specific hypotheses, based on such investigation, which have empirical import and are of primary interest to the nonlinguist.’ (Greenberg, 1963:74)

Ferguson and Joseph Greenberg. The results of this impressive project were published in 4 volumes in 1978. In the third volume *Word Structure*, there is an article ‘Typology and Universals of Personal Pronouns’ by David Ingram. Here, even more strongly than in the case of Forchheimer’s study, what is presented as results is assumption in disguise. This is even more startling as Ingram is explicit about his assumptions and questions. Still, he does not seem to notice that his conclusions are implied in his assumptions. Ingram is explicit about the questions he will address. The main question he asks is ‘what are the roles or combinations of roles in the speech act that each language considers to be of sufficient importance to mark by a separate lexical form?’ (Ingram, 1978:215-216). To investigate this question, Ingram chose to use the data as presented by Forchheimer in his book. Yet, one should be very careful in using Forchheimer’s data. In fact, Forchheimer’s data can only be used for the particular goal for which they were compiled: to make a morphological typology of number marking in the pronominal domain. Ingram does not realise that it is unsafe to formulate cross-linguistic conclusions about the paradigmatic structure of person marking on the basis of the data presented by Forchheimer. He simply takes Forchheimer’s data and computes the statistical frequencies of particular paradigmatic types. Forchheimer’s sample, however, was designed to show the different morphological possibilities to code number in the pronominal domain. Forchheimer’s data do not give an accurate statistical survey of the world’s linguistic possibilities of paradigmatic structure. Ingram’s conclusions thus stand on shaky ground.

‘Four systems ... are more frequent than others. ... The six person system [I, thou, he, we, you, they] is by far the most common system in languages.’ (Ingram, 1978:219)

More problematic still is Ingram’s proposal of universal features to analyse the systems found. He offers the features ‘speaker’, ‘hearer’ and ‘other’ for person and ‘singular’, ‘dual’, ‘trial’, ‘plural’ and ‘unmarked’ for number. These features are not results of the cross-linguistic study, but are explicit statements of the assumptions. Householder comments as follows on these ‘results’:

‘He [referring to Forchheimer, but it also applies to Ingram] chose to investigate the words which might conceivably translate *I, thou, we, you, he, she, it, they*. Could he possibly have taken any other step which would as surely guarantee in advance that he would discover nothing?’ (Householder, 1955:94)

In the late 1980’s Greenberg started a project entitled *The Diachronic Typology of Pronominal Systems*. Despite the name, the articles that he published during this project (Greenberg, 1985; 1988; 1993) deal with cross-linguistic variation, more than with typology or diachrony as such. He presents ample data showing the variability of linguistic structures to be found in the world’s languages, and hardly claims any restrictions on the possible linguistic structures. Increasingly through the articles, Greenberg is taken by the problem of the traditional terminology, thus questioning as such the assumptions implied. The rhetorical structure in these articles is completely different from the investigations I have discussed so far. In these articles, Greenberg considers a particular traditional grammatical term to be a theoretical proposal for linguistic analysis. He no longer takes these terms for granted (as they have been for centuries), but discusses whether they might be improved upon. He shows different instantiations of the terms in different languages and, finally, he asks whether the phenomena attested form a category. Instead of assuming implicitly the correctness of

the traditional terminology, he transforms the traditional terminology into a question. For example, in the article about the category ‘first person inclusive dual’ (Greenberg, 1988), he notes that the category to indicate ‘you and me’ (if it exists in a particular language) is ambiguous cross-linguistically. It is either a special category of person, or a category of number. He concludes:

‘Thus the first person inclusive dual appears to be an ambiguous category. ... We have sacrificed the notion of a uniform and universally valid set of typological categories by positing an ambiguous one.’ (Greenberg, 1988:12)

In another article, Greenberg questions the traditional three-person analysis (Greenberg, 1993). But although he shows that various linguistic phenomena do not fit in with the three-person analysis (such as zero third persons and inclusive persons), in the end he defends the three person analysis because it seems to occur so often among the world’s languages.

‘the traditional notion of a pronominal category of person with three members seems defensible. After all, the three-person pattern may be the result of historical accident, but it is a recurrent one.’ (Greenberg, 1993:20)

The approach taken by Greenberg in these later works will be followed in the present investigation. The traditional categories that are used to analyse person marking are not taken for granted, but tested on their accuracy against the available cross-linguistic variation.

1.4.4 Mühlhäusler and Harré

The starting point for Mühlhäusler and Harré in their work *Pronouns and People* (1990) is diametrically opposed to that of the authors discussed so far. The implicit, or sometimes even explicit, goal of the preceding authors was to find universals of human language, or at least restrictions in the attested linguistic structures. Mühlhäusler and Harré instead focus on the extent of linguistic relativity; they assume a tight interrelation between language and the social context and consequently stress the many ways in which linguistic structure can vary. They approach pronouns from their deictic side, leaving aside the anaphoric aspect of pronominal marking. To highlight the deictic approach, I would rather want to talk about ‘person’ instead of ‘pronoun’, but that is a minor terminological quibble.

‘A number of significant questions one might wish to ask about pronouns either have never been asked or else have received only marginal attention. Instead, linguists have chosen to address an issue, anaphoricity, that may be insoluble and that to us appears to be of only minor interest.’ (Mühlhäusler & Harré, 1990:59)

Mühlhäusler and Harré assume that person markers are found in small ‘closed sets’, reminiscent of what I call a ‘paradigm’. However, they seem to imply that each language has such a ‘closed set’, but I doubt that this is universally true (cf Thai, as discussed above in section 1.2.4). Still, the majority of the world’s languages indeed seems to have such a ‘closed set’ of person markers.

‘The term [pronoun] has been used in grammatical classifications of words to refer to a closed set of lexical items that, it is held, can substitute for a noun or a noun phrase ... When we talk of a ‘closed set’, we imply that in human language only a small, definite repertoire of pronoun forms is found in each.’ (Mühlhäusler & Harré, 1990:9)

Finally, Mühlhäusler and Harré stress the fact that the existing descriptive devices for the analysis of grammatical structure are in need of reconsideration and improvement, reminiscent of the approach of Greenberg in his later papers. Regrettably, Mühl-

häusler and Harré note the problem, but do not really put forward a new proposal for an improved cross-linguistic analysis of pronouns.

‘We thus need to have a descriptive apparatus sufficiently sensitive to describe all emic distinctions people actually make in all the world’s cultures in relation to the subject matter for which the etic concepts are designed, in our case the functioning of ... pronouns. As will be demonstrated eventually, the available etic tools may well be incomplete or even inadequate.’ (Mühlhäusler & Harré, 1990:61)

In general, the work by Mühlhäusler and Harré is a valuable addition to the cross-linguistic knowledge about the variability of person marking among the world’s languages. The main omission in their work, however, is the aspect opposite to variability: the restrictions on the variation. Mühlhäusler and Harré do not investigate the question of which of the many phenomena they discuss are common and which are rare. Neither do they present a geographical or structural analysis of the variation. Yet, they did not intend to cover these questions. Consequently, they leave those questions to be answered by others. I will attempt to do so to some extent in the present investigation.

1.4.5 Laycock on New Guinean pronouns

Laycock’s (1977) article on pronoun systems in Papuan languages is restricted in extent and scope. It deals only with an areally restricted set of languages, although from an area that is generally seen as one of the most linguistically diverse regions of the world: New-Guinea. Within this areal restriction, he only included independent pronouns in his comparison.³² He acknowledges that most languages seem to have more than one pronominal system, all of which can have a different internal structure.

‘A language may – and usually does – have a number of pronoun systems. Pronouns may be free or bound, emphatic or non-emphatic, full or abbreviated; they may also appear with different case-endings. Each such set may form a different system in that different categories are distinguished.’ (Laycock, 1977:33)

Other than the authors discussed previously, Laycock does not assume that there is one central pronoun system without indicating how to select that system out of the multiple possibilities. Laycock instead takes a formal characteristic to select his data: he compares those sets of person markers that are morphologically independent. By using this methodology, Laycock cannot draw any conclusions about the structure of a complete language. It could very well be that a particular opposition is not attested in the independent pronouns, but that this opposition can be found somewhere else in the language (viz in the inflectional person marking). Because of this method, Laycock cannot speak about the structure of language as a whole, but this drawback is compensated by the fact that he is immune to the criticism of Eurocentrism. If one searches for the presupposed categories known from ‘Standard Average European’ languages, one will probably find them somewhere in any language. The real question should be whether the phenomena found form a category in the language. This was one of Householder’s criticism of Forchheimer.

‘It is nonsense to talk of a category wherever a language ‘possesses the means for expressing a concept’ ... since, by definition, any natural language possesses the means

³² There are a few investigations of person marking that explicitly restrict the scope of the data to independent pronouns, cf Chlenova (1973), Sokolovskaya (1980) and Schwartz (1986).

for expressing anything. It is not the ability to express, but the inability to leave unexpressed that usually shows the presence of a category.' (Householder, 1955:94)

Laycock's main conclusion is that the pronominal systems of Papuan languages are not in conformity with Greenberg's Universal 42 ('all languages have pronominal categories involving at least three persons and two numbers', see section 1.4.3). He finds pronominal systems consisting of only two or three members (Laycock, 1977:35). I will specifically discuss this criticism of Greenberg's Universal 42 in section 3.6.2.

Laycock's approach –to include only independent pronouns in the comparison– is a methodologically coherent method, although it is restricted in its scope. I will extend this approach by including each and every person marking paradigm, whether independent or inflectionally marked (see section 1.2.3). The same pro's and con's arise. On the negative side, this method does not leave any room for conclusions about the structure of a language as a whole. Yet, on the positive side, it restricts the influence of Eurocentrism, which can otherwise be found in the form of implicit decisions to include the most familiar paradigm from a particular language in the investigation.

1.4.6 Conclusion

The focus and methodology of the present investigation have their clear precursors. First, the focus on person deixis is shared with Mühlhäusler & Harré (1990). Second, the interest for the paradigmatic structure is also found in the work of Laycock (1977) and Ingram (1978). The concern to find a suitable meta-language for the cross-linguistic variability of person deixis is shared with Mühlhäusler & Harré (1990) and the later work by Greenberg (1988, 1993). The wish to lay out the cross-linguistic diversity that is attested among the world's languages is also the goal of the work by Forchheimer (1953) and Mühlhäusler & Harré (1990). Finally, the attempt to formulate typological restrictions on this diversity is also found in the early work by Greenberg (1963) and in the work by Ingram (1978). In sum, these goals present an ambitious plan for the present work.

1.5 Outline of the book

This book is divided into three parts. The first part (chapters 2 and 3) deals with person marking. The second part (chapters 4, 5 and 6) investigates the paradigmatic structure of person marking. These two parts form the core of the present study. The third part (chapters 7, 8 and 9) adds an analysis of number marking to this core. The results of the various parts are summarised in chapter 10.

Part One, Person Marking, starts off with a cross-linguistic analysis of singular person marking (chapter 2). The three singular categories 'speaker', 'addressee' and 'other', which will be the basis for the investigation throughout this work, are introduced. A typology will be presented of paradigms that do not distinguish all three categories. Also, the status of zero marking of singular categories is investigated in detail. Building on this foundation, the marking of plural person marking will be investigated (chapter 3). The concept of 'plural' in the person marking domain is discarded and the notion 'group' is proposed as a replacement. The various theoretically possible group-categories are discussed against the background of the attested cross-linguistic varia-

tion. Also, a first advance towards a typology of the marking of these group-categories is made. A typology will be presented on the basis of the ‘first person complex’, which is the set of groups that include at least the speaker.

Part Two, *Paradigmatic Structure*, investigates the structure of the singular and the group categories combined. First, a complete typology of the attested paradigmatic structures is presented (chapter 4). This compilation is a lengthy and somewhat boring enumeration of case after case of paradigms of person marking. Along the way, a division into ‘common’ and ‘rare’ paradigmatic structures will be made. Some intermediate cases are classified as ‘semi-common’. On the basis of this qualitative typology, I have made a quantitative analysis of the structure of the paradigmatic structure (chapter 5). In this chapter, the focus will shift from the paradigms as a whole to the specific oppositions that are marked in each paradigm. Many implications between various oppositions are found, in the form of statements like ‘if a paradigm has opposition X, it will also have opposition Y’. Two major hierarchies are the result: the *Explicitness Hierarchy* and the *Horizontal Homophony Hierarchy*. Finally, these hierarchies are reformulated as a hypothesis for diachronic change (chapter 6). Two paradigms that are close to each other on these typological hierarchies are hypothesised to be similar. Diachronic change is taken to follow roughly the typological similarity space as described by the hierarchies. To test this hypothesis, cases of ‘cognate paradigms’ are collected. It turns out that the hypothesis fares rather well.

Part Three, *Number Incorporated*, extends the core of person categories as discussed so far with elements that explicitly mark number. Because I have replaced the traditional concept of ‘plural’ by the notion ‘group’ (in chapter 3), the traditional notion of ‘number’ has to be redefined within the person marking domain (chapter 7). Next, the typology of the core categories can be extended with these redefined number categories (chapter 8). The resulting classification of the various paradigms attested will be used for a quantitative analysis. These quantitative analyses result in a hierarchy that will be tested on its diachronic interpretation (chapter 9). The outcome will be a network of interconnected paradigms, which maps the similarity space of paradigmatic structure.

Finally, the main conclusions from these chapters are summarised in chapter 10. Also, some prospects for further research are presented there. When the main line of the present work is thus finished, a further theme is taken up in appendix A. There is a recurrent claim in the literature that there are languages with special person categories, called ‘compound pronouns’. In appendix A, I investigate these pronouns, and present the arguments why I disregarded these elements in the main body of this study. Appendix B and C present a survey of the typological classification as used for the quantitative analyses in chapters 5 and 8, respectively.

Part One

Person Marking

‘In those days his true self was still fighting with his assumed self, and winning. Person and persona, the man and his mask had separate identities then, he knew which was which.’

Jeanette Winterson, *Gut Symmetries*

The term ‘person’ is as old as the Western tradition of grammatical analysis. The English grammatical usage of the word ‘person’ is a loan-translation from the Latin grammatical term ‘persona’. In turn, the Latin *persona* was adopted from Greek linguistics, in which the equivalent word *prósopon* was used. This word is already attested in the oldest linguistic text in the occidental tradition that is known today, the *Tékhnē Grammatiké* by Dionysius Thrax (c. 100 BC).

‘Prósopa tría, prôton, deúteron, tríton; prôton mèn af hoû ho lógos, deúteron dè pròs hòn ho lógos, tríton dè peri hoû ho lógos.’ (Uhlig, 1883:51)

‘There are three persons, first, second and third. The first is the originator of the utterance, the second the person to whom it is addressed and the third the topic of the utterance.’ (Kemp, 1987:181)

The word *prósopon* was introduced comparing language metaphorically to a play. The Greek word *prósopon*, like the Latin *persona*, is the word for ‘mask’ in the theatre, hence for ‘dramatic character’ or ‘role’ (Lyons, 1977:638). The language user can take different roles in the linguistic interplay, changing his mask from speaker to addressee and back.

In this Part One, the various grammatical possibilities to mark person are investigated. Chapter 2 deals with the three basic concepts, ‘speaker’, ‘addressee’ and ‘other’. This division is almost identical to the classical analysis by Thrax. In this chapter, I will restrict myself to the singular usage of these concepts. Chapter 3 deals with the non-singular marking of person. The classical approach fares less well here. To be able to describe the cross-linguistic variation in non-singular marking, I will define the notion of ‘group’ as a replacement for the concept ‘plural’ in the domain of person marking.

Chapter 2

One among the crowd

The marking of singular participants

2.1 Introduction

Pronominal paradigms vary widely among the languages of the world. It will be a long journey to get this variation under control. I will proceed in small steps, limiting myself each time to a conveniently arranged subset of the variation. Inevitably, this will lead to some ad hoc decisions to include certain themes in one chapter or another, but everything that has to be said will come up eventually. This chapter deals with singular pronominal marking. This first step is a small step. In the introductory chapter, I have defined pronominal systems as paradigms that show at least an opposition between the marking for speaker and addressee. Both categories ‘speaker’ and ‘addressee’ refer to one singular participant. In this chapter, one category will be added to these two: the category ‘any other singular participant’. Many pronominal paradigms do not have one single morpheme for this category; the category is often split into different morphemes for different referents, like in the case of gender marking. However, some paradigms do have a single morpheme that is used for all ‘other’ singular referents. Only such paradigms will be considered in this chapter. Of these paradigms, only the singular morphemes will be considered. This may seem as if almost nothing is left to talk about, but even with such a restricted set of categories, there is enough variation among the world’s languages to fill a whole chapter. More important still, the foundation for the chapters to come will be laid with these three categories, which makes it necessary to be very precise at this point of the classification.

A proper definition of the terminology used is a basic aspect of all scientific inquiry. In a cross-linguistic comparison, this becomes even more prominent, as ‘exotic’ languages do not always behave as would be expected from a European point of view. In particular, some additions have to be made to the intuitive notion of ‘singular’ within the pronominal domain (section 2.2). When this groundwork has been finished, the principal objective of this chapter can be taken up: a typology of singular pronominal marking (section 2.3). Then follows a rather long discussion of pronominal paradigms with ‘impoverished’ singular marking. Various kinds of homophony between the singular categories are attested (section 2.4). I will argue that the cross-linguistic variation does not corroborate the hypothesis that the ‘richness’ of a pronominal paradigm is linked to the usage of independent pronouns (section 2.5). Subsequently, some at-

tention will be paid to the occurrence of zeros in singular pronominal marking. The question will be raised here whether these zeros should be considered as non-overt marking, or maybe better seen as non-existing marking (section 2.6). Finally, the content of this chapter will be summarised in section 2.7.

2.2 Definition of singular

From a West European point of view, singular pronominal marking may seem a straightforward (sub)category of linguistic structure, but the concept is in need of a more rigorous definition before it can successfully be applied to classify the world's linguistic variation. The first point to emphasise is that the use of the word 'singular' has to be taken literally: the forms that will show up in this section mark one *single* person or object. This may seem self-evident, but there are two problematic phenomena on the boundary between singular and plural marking that need some attention. One of these phenomena is the so-called 'inclusive', which is semantically plural, while some languages consider it structurally to be singular (cf Conklin, 1962). Notwithstanding such analyses, the inclusive marking will be disregarded in this chapter. The other problem is that the same pronominal elements are sometimes used for both singular and plural referents. Strictly speaking, such elements are not only 'singular', but they will still be included in this chapter as part of their meaning is singular. I will take the independent pronouns and the pronominal inflection from Axininca Campa, an Arawakan language from Peru, to exemplify these definitional considerations. The pronominal elements from Campa are shown in Table 2.1.¹

	<i>pronoun</i>	<i>prefix</i>	<i>suffix</i>
<i>speaker</i>	naaka	no-...	...-na
<i>addressee</i>	eeroka	pi-...	...-mi
<i>speaker & addressee</i>	aaka	a-...	...-ai
<i>other (masculine)</i>	irirori	ir-...	...-ri
<i>other (feminine)</i>	iroori	o-...	...-ro

Table 2.1: Campa pronominal marking (Reed & Payne, 1986:324-325)

The first problematic aspect for the definition of the term 'singular' is the special pronominal marker for the group of speaker and addressee combined, normally referred to as 'inclusive'. This inclusive is a special kind of pronominal marking that refers to a group that consists minimally of the principal speech-act participants, speaker and addressee, but more participants can be added. An inclusive is inherently plural as it minimally refers to two participants. However, in Campa, as well as in some other languages, the inclusive behaves morpho-syntactically like the singular persons. This can, for example, be concluded from the marking of plurality. Nouns and independent pronouns use the nominal plural suffix *...-payee(mi)*. This plural suffix can also be

¹ The different Campa languages are not mutually intelligible. However, the same 5-way pronominal paradigm is found all Campa languages. The data here come from the Apurucayali Ashéninka language.

found on the inclusive pronoun, as shown in (2.1 c). This indicates that the inclusive can be interpreted both as singular (only speaker and addressee) and as plural (speaker, addressee and others). For the plural interpretation, the plural suffix is needed (Reed & Payne, 1986:324).² In this chapter, the definition of the term ‘singular’ is taken to mean ‘exactly one’. Consequently, inclusive marking will not be discussed here, as it is not a form of singular marking. In the next chapter, I will return extensively to the inclusive.³

(2.1) CAMPA

- | | | |
|----|------------------------------------|---------------------------|
| a. | <i>t^{sh}imeri</i> | ‘bird’ |
| | <i>t^{sh}imeri-payeeni</i> | ‘birds’ |
| b. | <i>naaka</i> | ‘I’ |
| | <i>naak-payee</i> | ‘we (addressee excluded)’ |
| c. | <i>aaka</i> | ‘inclusive’ |
| | <i>aaka-payeeni</i> | ‘inclusive plural’ |

The second problem with the notion ‘singular’ is the occurrence of pronominal elements that cross the singular/plural boundary. The morphemes that are used for singular marking are, in some languages, also used when there is more than one person indicated. The same element can mean ‘I’, but sometimes also ‘we’. In such languages, number marking is often not obligatory, sometimes even non-existent. This is rather different from what is found normally in European languages, although the double meaning of the English pronoun *you* (which can be both singular and plural) is an example of this phenomenon. In Campa, the indeterminacy between singular and plural is taken much further than in English. There are no special pronominal affixes at all for plural marking in Campa. For example, the prefix *no-...* in its singular meaning ‘I’ is shown in (2.2).

(2.2) CAMPA

- no-kem-ake-ri*
 1-hear-TAM-3MASC
 ‘I heard him’ (Reed & Payne, 1986:325)

There are no separate element for meanings like ‘we’ or ‘they’ in the inflectional paradigms, so, strictly speaking, there is no number marking in these pronominal systems. There exists a way to mark plurality in Campa, by the (discontinuous) suffix *...-aiy-...-ni*, exemplified in (2.3). It remains undetermined what is pluralised by this suffix. Either, or both, of the participants in (2.3) could be pluralised by it. So, this (discontinuous) suffix is not a clear marker of plurality for the singular prefix *no-....*

(2.3) CAMPA

- no-kem-aiy-ake-ri-ni*
 1-hear-PLUR-TAM-3MASC-PLUR
 ‘We heard him’, ‘I heard them’, or ‘We heard them’ (Reed & Payne, 1986:325)

² The difference between the inclusive singular and the inclusive plural (also known as the augmented inclusive) will be discussed in sections 3.6.5 and 4.7.

³ In section 4.5.2, this type of paradigm as it is found in Campa, will be reconsidered. There it will be shown how it is possible that the ‘plural’ inclusive can sometimes fit in, structurally, with the singular marking.

Even in a looser sense, this plural suffix is no ‘real’ number marking, as it is in most cases not used when there is more than one person indicated. When it is clear from the context that it is a group of people to which the reference is made, then the plural marking is considered superfluous:

‘A notable phenomenon of Axininca person pronominals is that they are generally un-specific for number. The plural suffixes on verbs are ambiguous, except for context, as to plural subject or plural object. Furthermore, the person affixes often occur without the plural suffixes to indicate plural subjects of objects when it has been earlier specified in the discourse.’ (Payne, 1981:34)

In most cases, an utterance like (2.2), without plural marking, will be used to express statements like ‘we heard him’, ‘I heard them’ or even ‘we heard them’. The pronominal affixes in Campa do not specifically mark singular participants. In fact, they are unmarked as far as number is concerned. Such pronominal elements, that are unmarked for number, are included here under the heading of singular marking. For this chapter, the term ‘singular’ is taken to mean ‘minimally including singular reference’.⁴

Another definitional property of the subject of the present chapter is that subcategorisations are disregarded. In the example paradigms from Campa, there is a subcategorisation to gender in the third person; a distinction between the prefixes *ir-...* and *o-....*. It is an interesting question whether these two morphemes should be regarded as different subcategories of the third person or as two separate categories of pronominal marking. For now, I settle this discussion by simply disregarding all paradigms that do make these kind of distinctions, be it by gender or otherwise. In this chapter, I will discuss only those pronominal paradigms that do not have any subcategorisations in the singular. With this proviso, there is only a restricted (although still rather large) group of paradigms left that all have a ‘basic’ system for singular pronominal marking. In these ‘basic’ systems, there are maximally three different morphemes: one for speaker, one for addressee and one for any other. There can, of course, exist less than three different morphemes, like in the case of the English pronominal inflection, which has only two suffixes, *...-∅* and *...-s*.

To conclude, only those pronominal paradigms that show a morphological distinction between the traditional categories ‘speaker’, ‘addressee’ and ‘other’ are included in this section.⁵ The Latin inflectional pronominal paradigm is exemplary. There are three suffixes, *...-o*, *...-s* and *...-t*, and nothing else. Paradigms that have less than these three morphemes will be the crux of this chapter.

⁴ The distribution of forms that have both singular and plural reference will be taken up again in the next chapter. A complete analysis can be found under the heading ‘horizontal homophony’ in section 5.6.

⁵ I am tempted to use the abbreviations ‘S’, ‘A’ and ‘O’ for speaker, addressee and other. However, exactly the same letters are used to analyse ergative structures in the tradition of Dixon (1979). There, the letters ‘S’, ‘A’ and ‘O’ stand for subject, agent and object. To preclude confusion, I have decided to keep to the traditional numbers 1, 2 and 3 as abbreviations for speaker, hearer and other, respectively. The use of numbers should not be interpreted as a sign of inherent ranking of the categories; they are only used as abbreviations for reference to the basic speech act participants.

2.3 Typology

The three-morpheme paradigm, like the Latin singular suffixes, occurs frequently in the languages of the world. Still, there are cases where some of the three categories are coded by the same morpheme. I have found examples of pronominal paradigms with less than three singular morphemes in Germanic, among the non-Austronesian languages from New Guinea and in many other incidental cases spread out over the rest of the world. In such paradigms, different singular categories are marked by homophonous morphemes. In this section, a typology of these kinds of homophony will be developed.⁶

There are four theoretically possible kinds of homophony within the boundaries of the three persons. These four possibilities are shown in Table 2.2, along with the basic case, where all three persons are distinct. The five possibilities will be referred to as (Sa) through (Se); ‘S’ standing for singular. The capital letters in the table are variables designating identical morphemes by the same letter and different morphemes by different letters. Other combinations of letters than the ones presented do not yield any new possibilities. Type (Sb), for example, could just as well have been indicated by the letters ‘B’ for speaker, and ‘A’ for the combination of addressee and other. That would amount to the same paradigmatic structure. The present combinations of capital letters result from my decision to start always with ‘A’ at the top of the table. The Latin singular suffixes *...-o*, *...-s* and *...-t* form thus an (Sa) type paradigm. There are three different suffixes for the three categories. The Latin suffixes will be taken as a prototypical example of this first type (Sa). I will often refer to this type in the text as ‘Latin-type (Sa)’, or simply as ‘Latin-type’. Prototypical examples of the other types will be discussed later.

	(Sa)	(Sb)	(Sc)	(Sd)	(Se)
<i>speaker, ‘1’</i>	A	A	A	A	A
<i>addressee, ‘2’</i>	B	B	B		
<i>other, ‘3’</i>	C		A	B	

Table 2.2: Theoretically possible kinds of homophony

The first three possibilities in Table 2.2, (Sa), (Sb) and (Sc), have an opposition between speaker and addressee. Possibilities (Sd) and (Se), however, do not have this opposition. The opposition between the marking for speaker and for addressee is the defining characteristic of PERSON marking, and, consequently, (Sd) and (Se) do not mark person in the strict sense. Possibility (Sd) will be included in this section as a borderline case of person marking. It has at least an opposition between the speech-act dyad (speaker and addressee) and any other participant. However, possibility (Se) is

⁶ The word ‘homophonous’ is used as a purely descriptive label. It indicates that two categories, distinguished beforehand on independent grounds, are marked by the same morpheme. The frequency of certain kinds of homophony could be used as an argument against the distinguished categories. However, I do not have any theoretical presuppositions on which kinds of homophony would be expected, nor is there something *a priori* significant about certain kinds of homophony. The empirical results will point the way how to interpret the attested kinds of homophony.

excluded. With no person oppositions at all, it is not clear whether a paradigm really marks person, or rather something else. For example, consider the irrealis suffixes from Waskia, a Papuan language from New Guinea. The singular forms of these suffixes are shown in Table 2.3, just as they are presented in the grammar (Ross & Paol, 1978). The future and the imperative paradigms clearly mark person; they are Latin-type (Sa). However, the desiderative paradigm does not mark person. The suffix ...-ako only marks the desiderative mood; there is no distinction between the three person categories ‘speaker’, ‘addressee’ and ‘other’.⁷

	<i>future</i>	<i>imperative</i>	<i>desiderative</i>
<i>speaker, ‘1’</i>	...-iki	...-iko	...-ako
<i>addressee, ‘2’</i>	...-i	...-ko	...-ako
<i>other, ‘3’</i>	...-uki	...-uko	...-ako

Table 2.3: Waskia irrealis suffixes, singular only (Ross & Paol, 1978:68)

Pronominal paradigms of the Latin-type (Sa), with three different singular forms, are common among the world’s languages. This structure is well known and will not be exemplified further. I will come back to the Latin-type paradigm in section 2.6 to discuss the distribution of zeros in this kind of paradigms. The other three types – (Sb), (Sc) and (Sd) – are less common and less well known.⁸ In the next section I will discuss all examples of such paradigms that I know of. They turn out to be uncommon, but highly varied in their structure and widely dispersed among the languages of the world.

2.4 Singular homophony

2.4.1 Preamble

Some pronominal paradigms in the world’s languages do not distinguish all three singular categories – ‘speaker’, ‘addressee’ and ‘other’ – by different morphemes. Three different kinds of homophony will be considered in this section. The three kinds of homophony are summarised in Table 2.4. They will be discussed in turn. In the last

⁷ One could think of French inflectional person marking as an example of a homophony type (Se). There is no person marking in the singular (there is no difference in pronunciation; there is of course a difference in writing). However, there is an inflectional difference in the plural between the first and second person: *chantons* versus *chantez*. To my knowledge, this situation is highly unusual among the world’s languages. The only other examples known to me are the Icelandic Middle inflection (Thráinsson, 1994:162) and the past suffixes from the Papuan language Kapau (Oates & Oates, 1969:33-35). In all these cases there is an opposition between first person plural and other plural categories, but no opposition between the different singular categories. Notwithstanding these cases, I choose to hold on to the strict definition of person marking, which asks for an opposition between speaker and addressee, and consequently interpret the paradigms from French, Icelandic and Kapau as exceptional cases. Also note that the opposition *chantons* versus *chantez* does not clearly oppose speaker to addressee, as the addressee can be included in the reference of *chantons* (the ‘inclusive’ reading of the first person plural, see section 3.4).

⁸ For example, Noyer (1992:151-152) only discusses examples of the (Sd) homophony, and uses the existence of such cases to argue for a feature [\pm participant] in morphology. This argument becomes devoid of content in the light of the existence of the other kinds of homophony as shown in the next section.

column of the table, the type is labelled with the name of a well-known example of this structure. These names will be used to facilitate the reference to the various patterns.

<i>Type</i>	<i>description</i>		<i>mnemonic label</i>
(Sb):	speaker vs. the rest	first vs. non-first	‘Dutch-type’
(Sc):	addressee vs. the rest	second vs. non-second	‘Spanish-type’
(Sd):	other vs. the rest	third vs. non-third	‘English-type’

Table 2.4: Different kinds of singular homophony

2.4.2 Dutch-type homophony (Sb)

A homophony of type (Sb), ie an opposition between speaker and the rest, is, for example, found in the Dutch pronominal inflection.⁹ In the Dutch present, there is an opposition between a bare stem, which is used for the first person singular, shown in (2.4a), and a stem with a suffix *...-t*, which is used for second and third person, shown in (2.4b,c). As a mnemonic device, I will use the label ‘Dutch-type’ interchangeably for (Sb).¹⁰

(2.4) DUTCH

- a. *ik* *loop-∅*
 1SG,PRON walk-1SG
 ‘I walk’
- b. *jij* *loop-t*
 2SG,PRON walk-2/3SG
 ‘You walk’
- c. *hij/zij/het* *loop-t*
 3SG,PRON walk-2/3SG
 ‘He/she/it walks’

(own data)

In Dutch it is necessary to use an independent pronoun. A sole inflected verb stem is not a complete utterance in Dutch. It is often argued that an independent pronoun is necessarily added because of the structurally impoverished inflectional paradigm. However, this is not a necessary requirement for human language. There are also languages with the same Dutch-type inflectional paradigm, but without obligatory use of independent pronouns. In Lengua, a Mascoian language from Paraguay, the pronomi-

⁹ When I classify Dutch as (Sb), I refer to the inflection in Standard Dutch, with canonical word order subject-verb. The different Dutch dialects show a wide variety of homophony. All the different types distinguished in Table 2.2 are found within the dialectal variation of the Netherlands (van den Berg, 1949:7). Besides the dialectal variation, there is also variation within the Standard variety. The Standard Dutch inflection changes with the word order, a rather exotic phenomenon. In some contexts (like questions or when temporal adverbs are fronted for emphasis), the subject follows the finite verb. In these cases, the pronominal inflection is of type (Sd), with a zero first and second person, and a suffix *...-t* for the third person.

¹⁰ Within the Germanic family, the same Dutch-type homophony is found in the present inflection of the Icelandic ‘weak’ verbs. These verbs all have a zero first person singular, and identical marking for second and third person. The morphemes that are used for the combined second/third person differ, though, according to the class of the verb (Thráinsson, 1994:158-161).

nal prefixes show a homophony like Dutch, exemplified in (2.5). Unlike the Dutch situation, it is not necessary to add an independent pronoun in Lengua. It is left to pragmatic inference to identify the referent. Independent nouns (or independent pronouns) can, of course, be added, as shown in (2.5 c).

(2.5) LENGUA

- a. *ék-çl̄ĩngkyik*
1-go
'I go' (Susnik, 1977:98)
- b. *ab-l̄ĩngàé*
2/3-hear/feel
'You hear/feel', 'He/she/it hears/feels' (Susnik, 1977:99)
- c. *ab-waaktêyik* *sēnçlit*
2/3-return man,DEM
'That man returns' (Susnik, 1977:99)

The pronominal suffixes in Chitimacha, an extinct language from the USA, also show a structure of the Dutch-type (Sb), exemplified in (2.6). As in Lengua, it is not felt necessary to clarify the ambiguity of the 'non-first' suffix in Chitimacha:

'The [non-first] includes reference to either second or third person, and the actual reference in particular situations depends on the context. Ambiguity may be avoided by the use of the [independent] personal pronoun ..., but apparently the possibility of confusion is not as great as one might suppose, for sentences without independent pronouns are very common.' (Swadesh, 1946:324)

(2.6) CHITIMACHA

- a. *get-ik*
beat-1SG
'I beat'
- b. *get-i*
beat-2/3SG
'You beat', 'He/she/it beats' (Swadesh, 1946:317)

The Dutch-type homophony is relatively common among the Papuan languages from New Guinea (see, for example, the references in Haiman, 1980:xl). Among these languages, the morpheme for the 'non-first' is sometimes zero. This is the opposite markedness situation as in the case of Dutch, where the first person morpheme was found to be zero. Apparently, it is possible for either of the two elements in an (Sb)-paradigms to be zero. A zero non-first morpheme is found, for example, in the Papuan language Wambon, as shown in (2.7). In Wambon, just as in Lengua and Chitimacha, the use of independent pronouns is not obligatory. They can be used to disambiguate the reference of the 'impoverished' pronominal paradigm, but this is not necessary (see section 2.5).

(2.7) WAMBON

- a. *andet-ep-mbo*
eat-1SG-PAST
'I ate'

- b. *andet-Ø-mbo*
eat-2/3SG-PAST
'You/he/she/it ate' (de Vries, 1989:24)

To conclude this survey of the Dutch-type (Sb) homophony, I quickly review a few examples from around the globe. First, the pronominal prefixes for intransitive subject in the Siberian Chukotko-Kamchatkan languages (viz Chukchee, Koryak, and Kamchadal) show this structure (Comrie, 1980a). Second, a Dutch-type homophony is found in some of the Nakh-Dagestanian languages from the Caucasus. Helmbrecht (1996:136-138) presents the languages Tsakhur, Zakatal', Akhvakh and Megeb as showing an opposition between first and non-first. Both in the Chukotko-Kamchatkan and in the Nakh-Dagestanian case, the non-first categories are marked as zero. Finally, in Kenuzi-Dongola, a Sudanic language in Africa, there is a regular homophony between second and third person, through all tenses and aspects. The examples in (2.8) are from the Dongola dialect.

(2.8) KENUZI-DONGOLA

- a. *ai* *tóg-ri*
1SG,PRON beat-1SG
'I beat'
- b. *er* *tóg-im*
2SG,PRON beat-2/3SG
'You beat'
- c. *ter* *tóg-im*
3SG,PRON beat-2/3SG
'He/she/it beats' (Reinisch, 1879:67)

2.4.3 Spanish-type homophony (Sc)

In paradigms of type (Sc), the marking of speaker and other is identical, as opposed to the marking of addressee. In other words, there is an opposition between the marking of second person and the marking of non-second person. Seen in this way, (Sc) is a complementary structure to the Dutch-type (Sb) homophony, which had an opposition between first and non-first person. Examples of type (Sc) are rare, even compared to the scarcity of examples of (Sb). Still, paradigms of type (Sc) exist. Perhaps surprisingly, this type of homophony is found in the Spanish inflection. It is not found in all inflectional paradigms, but still in quite a few. It is exemplified in (2.9) with forms of the 'pretérito imperfecto'. For the rest of this section, I will use the label 'Spanish-type' as a mnemonic device, interchangeably with (Sc). In Spanish, this homophony is not only found in the pretérito imperfecto. In fact, it is found in many of the various tense-aspect-mood inflections. In Table 2.5, all non-periphrastic inflectional tense-aspect-mood categories are presented. Many are of type (Sc) (Hallebeek *et al.*, 1994:114-117).

(2.9) SPANISH

- a. *habla-ba*
speak-1/3SG,PAST
'I spoke', 'He/she/it spoke'

- b. *habla-bas*
 speak-2SG,PAST
 ‘You spoke’

(Hallebeek *et al.*, 1994:114)

The frequent occurrence of this homophony may be surprising as Spanish is often quoted as a prototypical example of a language with ‘rich agreement’ (eg Harbert, 1995:221-222). Rich agreement in turn is presented as one of the preconditions for pro-drop. Spanish shows indeed pro-drop, so the attested homophony between the first and third person would still qualify these paradigms as ‘rich’. Some doubts will be raised on this analysis in section 2.5.¹¹

<i>(Sa)</i> ‘Latin type’	<i>(Sc)</i> ‘Spanish type’
indicativo presente	indicativo pretérito imperfecto
indicativo pretérito indefinido	indicativo condicional
indicativo futuro	subjuntivo (all tenses)

Table 2.5: Spanish non-periphrastic tense-aspect-mood paradigms

The next example of a Spanish-type homophony is found in the Papuan language Koiari. In Koiari, the portmanteau suffixes for person-tense-mood display a typical Papuan realis-irrealis opposition. In the realis paradigms, the first person singular is in all tenses equivalent to the third person singular (using the suffix *...-nu* or *...-ma*) but different from the second person singular (using a suffix *...-nua* or *...-a*). This is shown in (2.10) for the past suffixes.¹²

(2.10) KOIARI

- a. *da* *ereva-nu*
 1,PRON see-1/3SG,PAST
 ‘I saw it’

- b. *a* *ereva-nua*
 2,PRON see-2SG,PAST
 ‘You saw it’

- c. *ahu* *ereva-nu*
 3,PRON see-1/3,PAST
 ‘He/she/it saw it’

(Dutton, 1996:24)

The defective pronominal paradigm in Koiari is supplemented by independent pronouns. These pronouns in Koiari are obligatorily used. This can be seen as all examples in the short grammatical sketch by Dutton (1996) have an overt subject, either by an independent pronoun or a full noun phrase. The use of an independent pronoun does not add emphasis. Instead, pronouns are fronted to add emphasis, and there is a special suffix *...-ike* used in those cases. When such a fronted emphatic pronoun is

¹¹ The other favourite example of a language with ‘rich agreement’ is Italian. Indeed, standard Italian does not show any homophony in the singular. However, Spanish-type homophony is found in some regional Italic languages. An example is the ‘imperfetto’ inflection from Siciliano (Bigalke, 1997:60-61).

¹² In the closely related language Ömie, a comparable non-second person homophony is attested, notably in the present tense (Austing & Upia, 1975:544).

used, the ‘unmarked’ independent pronoun still shows up in the sentence, as shown in (2.11). This indicates that the independent pronoun is obligatorily present.

(2.11) KOIARI

da-ike, *kailiki-ge* *da* *guramarero*
 1SG,PRON-EMPH PLACE-LOC 1SG,PRON sit¹³
 ‘I live at Kailiki’, ‘I am the one who is living at Kailiki’ (Dutton, 1996:64)

Another example of a Spanish-type homophony is found in Ika, a Chibchan language from Peru. Most morphology in Ika is suffixed, but the pronominal elements are prefixes. For the marking of the subject there is only one prefix in the singular: a second person *nΛ*-.... There is no overt marking for the other singular participants. This amounts to a homophony of type (Sc), with a zero for the non-second person marking.

	<i>Immediate Past</i>		<i>Past</i>		<i>Far Past</i>
<i>speaker ‘1’</i>	∅-tšua	uwin	∅-tšua	ukuin	∅-tšua-na-rua
<i>addressee ‘2’</i>	nΛ -tšua	ukuin	nΛ -tšua	užin	nΛ -tšua-na
<i>other ‘3’</i>	∅-tšua	Λwin	∅-tšua	užin	∅-tšua-na

Table 2.6: Ika singular person marking (Frank, 1985:89)

Just as in Koiari, the language Ika has a mechanism to disambiguate first person from third person. All ambiguity in Ika vanishes once the total marking in the sentence is taken into account. However, in Ika, ambiguity is not resolved by the use of independent pronouns, but by the special structure of auxiliaries. There is ample use of auxiliaries in Ika. The auxiliaries incorporate the marking of tense and evidentiality, among other aspects. Interestingly, time deixis interferes with person deixis in these auxiliaries. In Table 2.6, the singular forms of the various past tenses of the verb *tšua*, meaning ‘to see’, are shown. Note that the same form of the auxiliary, *ukuin*, is used for immediate past in the second person, but also for unmarked past in the first person. The use of this auxiliary thus implies some sort of person deixis. It is unclear whether these auxiliaries should be interpreted as tense markers or as person markers. They are something in between tense and person markers. The author of the grammar concludes:

‘Considered as markers of ‘degree of relevance’ or ‘distance between verbs and reference point’, the [auxiliaries] make more sense. Person, time and (un-)witness all enter into the relationship between an event and the speech situation.’ (Frank, 1985:90)

Note that the independent pronouns are normally not used (Frank, 1990:26). Most of the participant reference, besides the discussed inflectional marking, is zero in Ika discourse (Frank, 1990:121-122).

Finally, examples of the Spanish-type (Sc) homophony are also found in the Germanic family. Currently, it is found in the preterit inflection of ‘weak’ verbs in modern Icelandic. The morphophonological form of these suffixes in Icelandic falls apart in 6 classes, but all show the same structural (Sc) homophony (Thráinsson, 1994: 158-161). Despite this homophony, the pronouns are not used obligatorily in Icelandic

¹³ This verb is in an irrealis tense, which is not marked for person (Dutton, 1996:23).

(Thráinsson, 1994: 168-169). The Spanish-type homophony in the preterit is also found in older versions of Germanic languages, notably in Gothic, Middle Dutch (Schönfeld, 1959: 144-146) and in Old English (Robertson & Cassidy, 1954: 141). Interestingly, the Old English impoverished paradigm was not connected to the obligatory use of the subject pronoun. The pronoun could be dropped in Old English, although an overt subject reference should not be too far away:

‘In PDE [Present Day English] it is obligatory for the subject position ... to be filled ... In early OE [Old English], the NP was facultative, as in Latin ... However, in OE, ‘subjectless’ sentences were restricted to cases where the subject had been named in the preceding clause.’ (Görlach, 1997:91)

In all early Germanic languages, the independent pronouns were not as strictly obligatory as they are today, but they were still regularly used (Howe, 1996: 11-13).

The distribution of zeros among the presented examples of the Spanish-type homophony is asymmetric. All examples from the Germanic languages have a zero for non-second person and overt marking for the second person. The same distribution of zeros is also found in Ika. I have not found an example of the opposite distribution (non-second overt and second zero) among the languages of the world. As there are only a few examples of (Sc), it should not be concluded that it is impossible for a paradigm of type (Sc) to have a zero second person. Maybe, if more paradigms of this type are collected, examples will turn out to exist. Up till now, I know of none.

2.4.4 English-type homophony (Sd)

In paradigms of type (Sd), speaker and addressee are marked identically, as opposed to the marking of the category ‘other’. This paradigm is a borderline case of PERSON marking, as there is not opposition between speaker and addressee. English is one of the ‘exotic’ cases that show this homophony in the marking of singular participants. English has a homophony of type (Sd) in its inflectional marking: zero for first and second person, ...-s for third person. I will use the label ‘English-type’ as a mnemonic device for this paradigmatic structure of person marking, interchangeably with (Sd).

The same English-type homophony, but with opposite markedness as in English, is found in the present tense of the Caucasian language Hunzib, a Nakh-Dagestanian language from the Caucasus. The present tense suffix has two different forms: ...-č(o) for first or second person and zero for third person. Examples are shown in (2.12). The same homophony of speaker and addressee reference is also found in the related Nakh-Dagestanian language Lak (Helmbrecht, 1996: 131).

(2.12) HUNZIB

- | | | | |
|----|---------------------------|-------------------|-------------|
| a. | <i>də</i> | <i>hĩyaa-č</i> | <i>əcu</i> |
| | 1,PRON | open-1/2,PRES | door |
| | ‘I (shall) open the door’ | | |
| b. | <i>mə</i> | <i>bok’o.l-čə</i> | <i>heʔe</i> |
| | 2,PRON | gather-1/2,PRES | walnut |
| | ‘You will gather nuts’ | | |
| c. | <i>oʎul</i> | <i>hĩyaa-∅</i> | <i>əcu</i> |
| | DEM | open-3,PRES | door |
| | ‘He/she opens the door’ | | |

(van den Berg, 1995: 83)

This paradigmatic structure is also found in the imperfect tense of the South Caucasian language Svan, a neighbour of Hunzib, though unrelated to it. It is questionable whether this marking really should be considered pronominal marking. It is probably better to consider it as tense marking, with a slightly different status of the third person in the paradigm (Tuite, 1997:23-29). Just as in Ika, in the previous section, different deictic categories (time and participant deixis) are intermingled, making it difficult to classify the paradigm as one or the other.

A paradigm of the English-type with both third and non-third overtly marked, is found in the Papuan language Waskia. Waskia has a typical Papuan realis-irrealis distinction in the verbal inflection. The realis marking consists of three tenses (present, past habitual, past simple). All three have a English-type homophony. The forms of the present tense are shown in (2.13). Examples (2.13 a,b) show the suffix *...-sam*, that is used both with first and with second person. Example (2.13 c) shows the suffix *...-so* that is used with third person.

(2.13) WASKIA

- a. *ane itakta yu na-sam*
 1,PRON now water drink-1/2SG,PRES
 ‘I am drinking water now’ (Ross & Paol, 1978:67)
- b. *'ai ni ait omu arigi-sam i?*
 hey 2,PRON bird DEM see-1/2SG,PRES WH
 ‘Hey, do you see that bird?’ (Ross & Paol, 1978:112)
- c. *Gagi kaemkasik ko nagu-so*
 NAME evil spirit about fear-3SG,PRES
 ‘Gagi fears the evil spirit’ (Ross & Paol, 1978:49)

Independent pronouns are regularly used to clarify the subject, like in examples (2.13 a,b). The author of the grammar remarks that the subject is normally overtly present:

‘In unmarked sentences ... the subject is *not* normally omitted from declaratives and questions, but it is often deleted from imperatives. It is also missing in all impersonal sentences.’ (Ross & Paol, 1978:8, emphasis added, MC)

In the grammar (Ross & Paol, 1978), there are not enough examples of singular subjects to check this statement. However, a different picture emerges when all subject marking is counted in the short accompanying text (Ross & Paol, 1978:110-115). It turns out that, contrary to the above quote, it is rather normal to omit the subject. Relative to the 46 inflected verbs in the included text (counting all three persons, imperatives omitted), there are only 26 overt subjects (46%). Slightly more than half of the subjects are omitted in this text. This indicates that the presence of a subject is not really obligatory, although it is common.

Finally, I know of two examples of English-type homophony from American languages. The first example is the Oto-Manguean language Pame that has a suffix *...-p* for third person and a suffix *...-k* for non-third (Manrique, 1967:344). The second example is the Sahaptin language Nez Perce. In intransitive constructions, as shown in (2.14), the third person is marked by a prefix *hi-...* and the non-third by zero marking.

The independent pronouns can be optionally added; they are shown in brackets in the example sentences below.

(2.14) NEZ PERCE

a. ('iin) páayna
1,PRON arrive
'I arrived'

b. ('iim) páayna
2,PRON arrive
'You arrived'

c. ('ipi) hi-páayna
3,PRON 3-arrive
'He arrived'

(Rude, 1985:31)

2.4.5 Summary

This concludes the survey of languages that show some kind of homophony in the marking of singular participants. Table 2.7 summarises the cases presented. In the table, the three different types – (Sb), (Sc) and (Sd) – are subdivided as to the occurrence of zeros in one of the categories.

<i>Homophony Type</i>	<i>Occurrence of Zeros</i>	<i>Examples</i>
'Dutch-type' (Sb)	no zeros:	Lengua, Chitimacha, Kenuzi-Don-gola
	first is zero:	Dutch, Icelandic
	non-first is zero:	Wambon, Chukchee, Koryak, Kamchadal, Tsakhur, Akhvakh, Zakatal', Megeb
'Spanish-type' (Sc)	no zeros:	Spanish, Siciliano, Koiari, Ömie
	second is zero:	–
	non-second is zero:	Ika, Icelandic, Gothic, Old English, Middle Dutch
'English-type' (Sd)	no zeros:	Waskia, Lak, Nez Perce
	third is zero:	Hunzib, Svan
	non-third is zero:	English, Pame

Table 2.7: Typological summary of the examples

Only one of the resulting nine possibilities is not attested in this survey. Perhaps this last case will appear when the search is continued. I see no reason why this one pattern would not be possible, if all the others are.¹⁴

¹⁴ However, as it stands now, the one missing case from Table 2.7 precisely corresponds to a claim made by Uspensky (1972).

'If a zero expression occurs in the form of a certain person in the indicative mood, then, included in the meanings thus expressed (ie by a zero mark) we find the meaning of the 3rd person or that of the 1st person.' (Uspensky, 1972:68)

Homophony of singular pronominal marking is a relatively rare phenomenon, but it definitively is a possibility for a human language. The large number of examples that were presented in this section might suggest that it is common for a pronominal paradigm to have a homophony in the singular. However, it is not. First, only a restricted set of all the pronominal paradigms of each language shows the homophony. The 28 languages that are mentioned in Table 2.7 have a homophony somewhere in the language; it is not a characteristic of all pronominal paradigms in the language. Second, all other languages that I have studied do not show any singular homophony at all. Relative to the approximately 350 languages that are mentioned in the present work, the 28 cases with a singular homophony make up only 8%. However, even this figure is grossly overestimated, as I have included every example of a singular homophony that I have found, but skipped over numerous non-homophonous cases. My guess would be that a controlled sample would yield no more than 1% of languages with a singular homophony somewhere in the language. A homophony in the singular is a well-known phenomenon in the Germanic languages, but other than that it is only found in rather ‘exotic’ languages (see Table 2.8). With this in mind, the Germanic structure of inflectional person marking suddenly becomes rather ‘exotic’.

<i>Europe:</i>	English, Old English, Icelandic, Gothic, Dutch, Middle Dutch, Spanish, Siciliano
<i>Caucasus</i>	Hunzib, Lak, Tsakhur, Akhvakh, Zakatal’, Megeb, Svan
<i>Siberia:</i>	Chukchee, Koryak, Kamchadal
<i>New Guinea:</i>	Wambon, Koiari, Ömie, Waskia
<i>America:</i>	Nez Perce, Chitimacha, Pame, Lengua, Ika
<i>Africa:</i>	Kenuzi-Dongola

Table 2.8: Geographic stratification of the examples

For the Germanic languages, it is known that historically the ‘impoverished’ singular marking arose because an original three way distinction of the Latin-type wore off through time. For most of the other examples that have been discussed in this section, there are no historical records. Moreover, often not enough data for a comparative reconstruction are available. It is thus impossible to say anything definitive about the way these singular paradigms arose. However, it seems perfectly possible that ‘impoverished’ paradigms can arise in other ways than through wear. Constructively, a

[continued from previous page]

The data that led to this claim are not made explicit by Uspensky, but he seems to have formulated his claim in exactly these words to account for the various inflectional structures of the Germanic languages. There is no sign in the short 1972 article that he knew of all the different possibilities that I have sketched in this section. Still, the formulation turns out to be a masterful generalisation over the cases that I have found. The only problem is that I can see no reason why this one case would be impossible for a human language. Also, there are a few counterexamples to Uspensky’s claim. They will be presented in the section 2.6.2. These counterexamples are Latin-type paradigms with a zero second person. This means that the counterexamples have different (overt) forms for first and third person reference (different from the missing Spanish-type (Sc) paradigm, that would have had the same form for first and third person).

morpheme for ‘speaker’ might arise alone, leaving the rest unmarked. In the same way, a morpheme for ‘addressee’ might arise without counterpart. Another path towards a homophonous paradigm might lead through mixed marking of different deictic categories, as in Ika and Svan. There may be a difference between the kinds of homophony that arise by wear and kinds of homophony that can arise constructively. However, more in-depth studies of languages with singular homophony are needed to be able to approach such a hypothesis.

An important generalisation over the examples that were presented in this section is that the different examples of singular homophony are all found in inflectional paradigms. I have not come across a homophony in a paradigm of morphologically independent pronouns.¹⁵ This generalisation can be formulated as an implication:

(2.15) Homophony Implication

singular homophony → inflectional paradigm

This implication can be reversed: independent pronouns never have a homophony in the singular. For independent pronouns it is necessary to clearly distinguish the three main singular categories. This amounts to the following implication:

(2.16) Homophony Implication (reversed formulation)

independent pronouns → speaker ≠ addressee ≠ other

It should not be concluded from this implication that all languages have at least three different overt independent pronouns – one for speaker, one for addressee and one for other. This conclusion is not allowed for two reasons. First, there are languages that do not have ‘real’ independent pronouns (see the introductory chapter). For those languages, the antecedent of the universal is not met. Second, the ‘other’ category is not as straightforward as the ‘speaker’ and the ‘addressee’ categories. In many languages, the ‘other’ category is not marked by a specialised pronominal morpheme. All that is implied in the statement in (2.16) is that the marking for ‘other’ is not identical to the marking of ‘speaker’ or ‘addressee’. The implication makes no claim for the specific linguistic form of the marking of the singular categories; they could also be zero. More specific claims about the form of paradigms that differentiate all three singular categories will be discussed in section 2.6. However, first I will discuss another aspect of singular homophony: the consequences for the structure of the rest of a language when there is singular homophony. The conclusion will be that languages do not treat paradigms with a singular homophony necessarily as an impoverished form of person marking. Although such paradigms mark less person distinctions, some marking is still considered better than none. In the following section, I will argue that structurally ‘impoverished’ paradigms with singular homophony are not necessarily functionally inferior to paradigms without singular homophony.

¹⁵ Noyer (1992:151, citing Lipkind 1945) mentions the Siouan language Winnebago as an example with a singular homophony in the independent pronouns. The pronoun *nee* is used for reference to the speaker and the addressee; the pronoun *?ee* is used for reference to any other participants. I have been unable to get hold of the original source, nor of any other primary source on Winnebago to check this claim.

2.5 Singular homophony and *pro-drop*

In some languages, a singular homophony is seen as a problem. The contemporary West Germanic languages, but also Koiari, Waskia and Ika, show a tendency to disambiguate the homophonous singular persons in some way. In Germanic and in Koiari, the disambiguation is achieved by using independent pronouns that are co-referential with the homophonous inflection. In Ika, a more intricate combination of various linguistic devices do the same job without the use of the independent pronoun (see section 2.4.3). For other languages, in contrast, it is noted explicitly in the descriptions that the singular homophony in the pronominal paradigm does not lead to a surplus of other markings. In Wambon, Lengua, Chitimacha, Spanish, Icelandic, Old English and Nez Perce, it is no problem that the singular homophony creates some ambiguity in the structure of the language. Most of the time, enough information is present in the discourse to fill the referential gaps that are left by a homophonous paradigm. The speakers of many languages do not find it necessary to disambiguate the three singular persons all the time. Seemingly, it is perfectly possible for a human language to do without a morphological distinction between all three persons.

To exemplify the different approaches to singular homophony, I will make a short comparison of the person marking in Wambon and English. Wambon and English are both languages with a singular homophony. In the English inflection, the speaker and the addressee are marked identically. In the Wambon inflection, the addressee and the other are marked identically (see page 42). Still, there is an important difference between these languages. English uses its independent pronouns obligatorily; Wambon does not. To show the differences and correspondences between the person marking in Wambon and English, I have analysed the use of the inflection and the independent pronouns as they occur in the Wambon texts in the grammar by de Vries (1989). I will concentrate on the marking of first person subjects, and compare the Wambon structure with the structure of the English translation by de Vries. This kind of comparison is full of pitfalls as the structure of Wambon differs strikingly from English. For example, Wambon makes ample use of switch-reference and serialisation. Both phenomena strongly influence the structure of the sentence. Still, I think that a comparison of the Wambon texts with the English translation is possible. Remarkably enough, the English translation consists of syntactically correct sentences of English while the style is rather poor. This poor style probably reflects the intention of the grammarian to translate the Wambon text as closely as possible to the original. I interpret the translation as being the closest possible reflection of the original stories in syntactically correct English. From this assumption it is possible to compare the way pronominals are used in the structure of both languages. The results of the counts are presented in Table 2.9.¹⁶

¹⁶ The counts are based on two texts from the grammar by de Vries (1989). The two texts ('the pig hunt', p. 117-118, and 'Sawing', p. 120-128) are chosen because they have numerous first person subjects. I counted first persons singular and plural for these texts. In the plural, Wambon also has a homophony with an opposition between first and non-first person, so the markedness situation is comparable to the singular. If I had only counted the singular pronominal elements, the number of instances would have been much smaller, and the results less revealing.

	<i>Wambon</i>	<i>English</i>
<i>Finite Verbs</i>	74	92
<i>Independent Pronoun</i>	9	80
<i>Pronouns per Finite Verbs</i>	12%	87%

Table 2.9: First person marking compared between Wambon and English

For the original Wambon texts, as well as for the English translation, the number of independent pronouns in subject function has been counted. The independent pronouns of Wambon hardly occur (only 9 cases). By contrast, independent pronouns are often used in the English translation (80 cases). Then, the number of finite verbs that are marked for first person subject is counted for Wambon (74 cases). For English, this count is not so straightforward, as there is no inflectional first person marking. In the English translation, I counted the number of finite verbs that have a first person subject (92 cases).¹⁷ The percentage of pronouns per finite verbs gives a rough indication of the differences in use between the independent pronouns of both languages. Wambon uses a pronoun for 12% of the finite verbs, against 87% for English. Also, the sparse use of the Wambon independent pronoun is not determined by a regular syntactic structure. The function of the independent pronoun of Wambon is to put emphasis on the subject.¹⁸ On the other hand, the number of independent pronouns in English (80 cases) roughly matches the number of person-marked finite verbs in Wambon (74 cases). From these counts, I tentatively conclude that the Wambon pronominal inflection functions roughly like the English independent pronouns. The Wambon inflection is the central person marking device in the language, not the independent pronouns. The person inflection in Wambon has a homophony of second and third person. Even though this language has an ‘impoverished’ person-marking system, it is not necessary to use the independent pronouns for disambiguation. English also has a singular homophony, and in this cases it is necessary to use the independent pronouns. This shows that the often-heard argument that English obligatorily uses its independent pronouns *because* of the impaired inflectional system does not hold as a universal characteristic of human language. As can be seen from the Wambon case, it is also possible for a human language to have an ‘impaired’ inflectional pronominal paradigm, but still not to use the independent pronouns obligatorily.

This conclusion is at odds with the ‘*pro*-drop’ hypothesis or, at least, it restricts the space for its interpretation. *Pro*-drop, the hypothesis states, is licensed by ‘rich-agreement’.

‘It has been long noted that *pro*-drop tends to occur in languages with rich subject-agreement morphology, such as Spanish and Italian, but not in languages such as

¹⁷ The number of English finite verbs (92 cases) is higher than the number of independent pronouns (80 cases) due to conjunction reduction (‘equi-NP deletion’). A sentence like ‘I pursued it and saw it lying close’ has only one first person subject pronoun (‘I’) while there are two finite verbs that have a first person subject (‘pursued’ and ‘saw’).

¹⁸ To make a better argument, the use of the independent pronouns should of course be investigated from the structure of the discourse. A full analysis of Wambon-discourse is beyond the scope of the present work. Impressionistically, the independent pronouns in Wambon are used at turning points of the story. This indicates that the use of the independent pronouns is at least partially regulated by the structure of the discourse.

French and English, which have relatively impoverished agreement morphology.’
(Harbert, 1995:222)

To my knowledge, it has never been made explicit what it means exactly for a personal agreement paradigm to be ‘rich’. However, it can be concluded from the examples in this chapter that it is possible to conflate the reference between the three basic singular categories and still have *pro*-drop. Thus, singular homophony does not necessarily make the paradigms less ‘rich’. The proposal that *pro*-drop is licensed by the structure of the inflectional pronominal paradigm is not unequivocally true. An equivalent singular structure to the one that licences *pro*-drop in Wambon (and other languages) prompts English (and other languages) to prevent *pro*-drop. To conclude, the ‘richness’ of a pronominal paradigm does not reside in the marking of singular participants. Maybe the richness of a pronominal paradigm can be found in the non-singular forms. This possibility will further be investigated in the following chapters. However, I will conclude eventually that singular homophony is the least explicit kind of paradigmatic structure (see especially section 5.5). This conclusion makes it highly improbable that the structure of a pronominal paradigm correlates cross-linguistically with ‘*pro*-drop’ characteristics.

2.6 Zeros in paradigms without homophony

2.6.1 Preamble

In the vast majority of pronominal paradigms in the world’s languages, there is no singular homophony in the paradigm; all three singular persons are marked by different morphemes. In such paradigms, each of the three participants (speaker, addressee and other) could, in principle, be zero. However, it turns out that not all participants are equally likely to be zero. It is very unusual to find paradigms with a zero for speaker or addressee reference. The few examples that I have found are discussed in section 2.6.2. In contrast, it is easy to come up with numerous examples of zeros for ‘other’ reference. Examples of third person zeros will be discussed in section 2.6.3. Moreover, for many of the cases with a zero third person, it is questionable whether this zero really exists. It might be better to consider a zero as non-existing in such cases. This argument will be presented in section 2.6.4.

2.6.2 Zero speaker or addressee

Zero marking for speaker or addressee reference is only scantily found among the world’s languages (Koch, 1994:31-37; Siewierska & Bakker, 2000). Many of the examples that appear, at first glance, to have zero marking for speaker or addressee, turn out to be rather poor cases. On closer inspection, either different descriptions of the same language do not agree on the analysis, or the zero turns out to be only one of the possible allomorphs for the first or second person marking.¹⁹

¹⁹ Questionable zeros were attested in the descriptions of Wiyot, Auyana, Maba, Rumanian and Marghi. In the Algic language Wiyot, a zero described for speaker reference by Reichard (1925:74), but a later description only lists zero as a possible allomorph for a suffix ...-a (Teeter, 1964:71). There is no analysis of the factors that govern this allomorphic alternation. In the Papuan language Auyana, a zero is attested for addressee reference. However, the zero alternates with a suffix ...-a, but it is not clear under which conditions (Bee, 1973:188). A comparable situation is found in the Nilo-Saharan language Maba. Tucker (1966:202) present paradigms with a zero second person, but it turns out later

First person zero marking is rare. I have not come across any clear example.²⁰ However, zero marked first persons occurs in the marking of kin-term possession (Croft, 1990: 145-146; Koch, 1994: 51-54). In the context of kin-term possession, the first person seems to be the unmarked participant and, consequently, it is sometimes found marked as zero. I can not judge whether the preference for zero first person marking is far-reaching in such constructions, because I have not included the marking of pronominal possession in the present investigation. Next, I turn to pronominal paradigms with a zero addressee reference. Such cases are slightly more frequent compared to paradigms with a first person zero, although the overall number of examples is so low that it is difficult to judge whether this difference is typologically valid. The first example with a zero addressee comes from Bongo, a Nilo-Saharan language from Sudan. In Bongo, the pronouns cliticise in front of the verb in the past tense, as in (2.17a), but remain independent when there is another word in between the pronoun and the verb, as in (2.17b). The first person singular pronoun *ma* cliticises as *m-*..., see (2.17c), but the second person singular pronoun *i* disappears altogether, leaving a zero, shown in (2.17d).²¹

(2.17) BONGO

- a. *b-ony*
3SG-eat
'He/she/it ate' (Santandrea, 1963: 62)
- b. *'da bu dukpa amony 'bu kururi kpáu*
then 3SG,PRON after eat eggs python all
'So he then ate all the eggs of the python' (Santandrea, 1963: 62, 94)

[continued from previous page]

that this zero is only one of the possible allomorphs (Tucker & Bryan, 1966: 195). Also in Rumanian, a zero first person is found, but only in the a-conjugation (Posner, 1996: 43), see also page 108 of the present work. Finally, Burquest (1986: 77) lists a zero for the speaker reference in the aorist of the Chadic language Marghi. However, this is probably a misinterpretation of the data in the original source. Burquest based his data on the description by Hoffmann (1963: 174). Hoffmann gives a paradigm in the aorist without first persons, but this is probably because the examples in the paradigm are questions. Questions to the speaker (ie yourself) are probably left out because they are pragmatically strange. There are other examples of aorists in the texts as presented by Hoffmann. In these examples, speaker reference is marked overtly.

²⁰ The only predicatively used pronominal paradigm with a zero marker for speaker reference, as far as I know, is found in the pronominal inflection in Alagwa and Burunge, two Southern Cushitic languages from Tanzania. All Cushitic languages mark gender in their pronominal inflection, so these paradigms do not count as 'pure' person marking following the definitions as presented in section 2.2. The gender marking complicates the paradigmatic structure as there is a regular homophony between the second person and the third person feminine (a common feature in Semitic languages). However, because these are the only example of a zero for the marking of speaker reference, I still want to mention the two languages here. It may be clear that it is not the most obvious and clear example. Data on Alagwa can be found in description by Whiteley (1958: 32-33) and Mous (in prep: 54-55). On Burunge, see Kießling (1994: 124). On first view, there is no zero first person in the closely related language Iraqw (Elderkin, 1988: 81-82). Still, after some cumbersome morphophonological analyses, Elderkin (1988: 87) concludes that 'verbs with obstruent stems seem to have \emptyset marking first person.'

²¹ A comparable phenomenon is found in the related language Bagirmi. For Bagirmi, a disappearing second person pronoun clitic is described in Gaden (1909: 10, 15). In a short description of another related language, Beli, it is noted that there are no clitic forms of the pronoun (Santandrea, 1963: 110).

c. *m-ony*
 1SG-eat
 ‘I ate’ (Santandrea, 1963:62)

d. *∅-ony*
 2SG-eat
 ‘You ate’ (Santandrea, 1963:62)

Another example of a zero marked addressee reference is found in Grebo, a Kru language from Liberia. Grebo, like all Kru languages, has independent pronominal elements. There is no pronominal inflection. In many Kru languages, the first and second person independent pronouns can be omitted when they come directly in front of certain auxiliaries; in those cases the person distinction is marked by different tones on the auxiliary (Marchese, 1978: 121). Grebo is a special case among the Kru languages. In Grebo, the pronouns can be omitted with all indicative verbs (also without the presence of an auxiliary). The bare verbs also show tonal differences to mark person (Innes, 1966:62-63). However, when the pronouns are overtly present (so there is no person marking tone), then there is no overt marking of the second person in the indicative mood (Innes, 1966:50,61). The second person exist in other moods (and in the indicative of other Kru languages, Marchese, 1978: 117-118) but not in the indicative of Grebo, as exemplified in (2.18).

(2.18) GREBO

a. *ne₁ du₁ ne₄ ne₃*
 1,PRON pound
 ‘I have pounded it’

b. *∅ du₄ ne₄ ne₃*
 2,PRON pound
 ‘You have pounded it’ (Innes, 1966:61)

Another kind of zero addressee marking is linked to the imperative. Imperatives are prototypically directed to the addressee; addressee reference is the unmarked situation for an imperative. Consequently, imperatives are normally zero marked for addressee reference (Kurylowicz, 1964:241; Greenberg, 1966:44; Koch, 1994:57-59). Imperatives are a possible source for zero marked addressees in other constructions. However, in most languages, the imperative is a special construction, not marked for person at all. In some languages, the imperative is generalised to include other ‘manipulative speech-acts’ (Givón, 1990:806-824). In such cases, this mood-variant of the predicate is often marked for person. This is, for example, found in the Papuan language Daga. The imperative in Daga has a ‘debitive’ meaning: ‘the imperative suffixes cover the area of meaning of ‘must’ and ‘should’ (Murane, 1974:56). This meaning is also relevant with first and third person subjects. The person marking of this debitative in Daga, presented in (2.19), shows a Latin-type paradigm with a zero addressee.

(2.19) DAGA

a. *war-ap*
 get-1SG,DEB
 ‘I must get (it).’

b. *wat-∅*
 get-2SG,DEB
 ‘You must get (it)’, ‘Get (it)!’

c. *war-ep*
 get-3SG,DEB
 ‘He must get (it)’

(Murane, 1974:56)

A comparable case with zero addressee reference is found in the irrealis suffixes in the Papuan language Waskia. These suffixes were shown earlier in this chapter in Table 2.3 on page 40. They can be analysed as showing a suffix ...-*i* for future and a suffix ...-*ko* for imperative. In both cases, the addressee marking is zero. Just as in Daga, this imperative in Waskia is better analysed as a debitive:

‘The imperative ... denotes not only the imperative as in English ... but also concepts expressed in the English modal verbs ‘must’, ‘should’ and ‘ought to.’ (Ross & Paol, 1978:69)

To conclude, a second person zero is not common, but it is certainly possible for a human language. However, when all three singular persons are marked differently, the speaker and addressee reference will almost universally be overtly marked. In contrast, the ‘other’ reference is often marked zero. This kind of zero marking will be discussed next.

2.6.3 Zero other

Zero marking for speaker or addressee is rare among the languages of the world. In contrast, a zero for the marking of the ‘other’ participant is found commonly, in all parts of the world.²² To conclude that a language has a zero is principally a dangerous statement. A zero is not directly observable, and there should be good reasons to propose an empty category. This is a general problem in all morphological analysis, but it becomes particularly important in the case of a zero third person. In contrast to the first and second person, the third person is not a positively defined category: a third person is defined negatively as the category that is not first or second. A good argument is necessary to propose zero marking for this ‘non-person’ (Benveniste, 1966:228-231). A reasonable argument for the existence of a zero third can be made when the first and second person marking are overt inflections. When an uninflected stem occurs that is paradigmatically equivalent to first and second person, and that is functionally in complementary distribution with first and second person, then it is fairly safe to propose a zero third person. I will now present some cases of zero third persons in inflectional pronominal marking. Afterwards, in section 2.6.4, I will devote some attention to third person zeros in independent pronominal marking, and discuss the question whether zeros really exist in those cases.

²² Hagège (1982:96) found in his sample that only 24% of the world’s languages have a ‘real’ third person pronoun. He does not present the data on which he bases this claim, except for a footnote that sounds impressive, but remains rather vague:

‘en ce qui concerne le corpus qui sert de base au présent ouvrage, on en a élargi le plus possible les dimensions: l’échantillonnage est constitué de 754 langues (également réparties sur les cinq continents), décrites dans les fiches personnelles de l’auteur, et dans d’autres sources: Trubetzkoy, Hagège-Haudricourt, Greenberg-Ferguson-Moravcik.’ (Hagège, 1982:12)

The Mongolian language Buriat is an example of a language with zero third person for subject reference, shown in (2.20). Note that the first and second person are suffixes, different from each other, and that the third person does not have a suffix. The verbs without suffix have a different reference. These verbs do not refer to speaker or addressee. This is enough reason to propose a zero third person suffix.

(2.20) BURIAT

a. *jaba-na-b*
go-PRES-1SG
'I go'

b. *jaba-na-š*
go-PRES-2SG
'You go'

c. *jaba-na-∅*
go-PRES-3SG
'He/she/it goes'

(Pope, 1960:57)

This phenomenon is found regularly in the languages of the world. The following list presents a rough outline of the geographical distribution. This is not intended as an exhaustive listing of this phenomenon, but it should give an impression of the areas where inflectional third person zero forms are found. To start, it is rather unusual to find zero third persons in Africa. There are only some incidental cases within Nilo-Saharan.²³ Travelling through the Eurasian continent, zero third persons are rarely found within Indo-European. Only the languages of the Brythonic subgroup of the Celtic languages have a zero third person.²⁴ It is found more often in Uralic, both in the Ugric languages and in the Samoyedic languages.²⁵ Also, zero third persons seem to be widespread in Altaic.²⁶ In contrast, zero third persons are only rarely found in the Semitic family.²⁷ A zero third person has been reconstructed for proto-Sino-Tibetan, and still a large part of the languages actually show a third person zero (DeLancey, 1987:806-809). Because there is hardly any bound pronominal marking among the languages from mainland Southeast Asia, there is also no chance to find a

²³ Eg Madi (Crazzolara, 1960:72-73,83-84), Temein (Tucker & Bryan, 1966:258-259), Didinga (Odden, 1983:154-157,166), and Tubu – a dialect of Teda (Lukas, 1953:111).

²⁴ Eg Breton (Stephens, 1993:373-375) and Cornish (George, 1993:446-447).

²⁵ In Ugric: eg Hungarian (Tompa, 1968:167), Mansi (Keresztes, 1998:399-400) and Khanty (Abondolo, 1998a:372-374). In Samoyedic: eg Nganasan (Helimski, 1998:502-504) and Nenets (Salminen, 1998:533-534).

²⁶ Eg Buriat, shown in example (2.20) and Turkish (Lewis, 1967:106).

²⁷ In roughly all Semitic languages/dialects the third person masculine of the perfect is zero. I do not include this as a case of third person zero in this list because there is generally an overt third person feminine suffix in the same paradigm. That is, I follow the general analysis of occidental linguists to put the feminine suffix in the same paradigm as the other pronominal suffixes, and, consequently, the third person is not completely zero. Note though, that the classical Arabic linguistic tradition did not include the feminine suffix among the pronominal suffixes, for rather good reasons (Versteegh, 1997:80-81). There is a real zero third person found in some of the modern Aramaic variants, in the enclitic personal pronoun. This enclitic personal pronoun does not have an overt third person in Mandaic (Macuch, 1965:155) and Neo-Aramaic Assyrian (Tsereteli, 1978:57,60).

'The classical imperfect died out, being replaced by the special Neo-Aramaic present, consisting of the particle q, qa, qi followed by the active participle with the enclitic personal pronoun.'
(Macuch, 1965:256)

zero third person on the basis of inflectional patterns. The same holds for many Austronesian languages, which have mainly independent pronominal marking. Impressionistically, it seems that zero third person forms are rare among the few languages with inflectional person marking. In contrast, third person zeros are common among the non-Austronesian languages in New Guinea. It is interesting that they are found regularly in paradigms for inflectional undergoer marking.²⁸ In Australia, roughly half of the languages have pronominal clitics. These show regularly third person zeros.²⁹ In North America, third person zeros are common; in South America they seem to occur much less.³⁰

2.6.4 Non-existing other

The zeros in the inflectional morphology, as discussed in the previous section, are rather straightforward. However, there are also languages that seem to have a zero third person marker in the paradigm of morphologically independent pronouns. These ‘missing’ markers are much more prone to be misinterpreted when they are called ‘zeros’. Often, a third person independent pronoun simply does not exist. The function that is fulfilled by an overt third person pronoun as attested in West European languages, is taken care of by other linguistic means in such languages. There are numerous possibilities for a language to express the function of an independent third person pronoun. For example, a demonstrative (‘this’, ‘that’ etc.) can be used, or a full noun phrase (‘the man’), or a proper name (‘John’). As every language already has multiple means to express third person, it seems better to assume *a priori* that non-bound third person pronouns do not exist. Such an approach seems better than to generate needless zeros.³¹

The kind of argument that shows that ‘zero’ third person independent pronouns are sometimes better interpreted as ‘non-existing’ is exemplified here with the pronominal marking from Yidjñ, a Pama-Nyungan language from Australia. Yidjñ has independent first and second person pronouns and no pronominal inflection or clitics on the predicate, as is shown in (2.21 a,b). The third person pronoun is ‘zero’. Normally, no marking at all is present for a third person subject, as shown in (2.21 c). Other linguistic means, like demonstratives or full noun phrases, can be used for clarification, but often these are not used because the reference is clear from the context. Moreover, there is positive evidence that the absence of pronominal marking does not imply the presence of a zero. In Yidjñ, the meaning ‘third person’ cannot be inferred from the

²⁸ Eg Daga (Murane, 1974:44), Yagaría (Renck, 1975:21) and Siroi (Wells, 1979:28).

²⁹ Dixon (1980:363-370) gives Western Desert languages and Walmatjari as examples.

³⁰ In North America, zero third person are found consistently in, for example, the Yuman languages (eg Maricopa, Gordon, 1986:15-21) and the Muskogean languages (eg Chickasaw, Payne, 1982:359). Other cases are found in Tonkawa (Hoyer, 1933:72-76), and Kwakiutl (Boas, 1947:252). In South America, examples are much more difficult to find. Incidental examples are the object affixes in Ika (Frank, 1990:52) and the object affixes in Quechua, for example in the Huallaga Quechua variant (Weber, 1986:334).

³¹ The same problem (whether non-overt independent pronouns are zero or simply do not exist) arises by the zero second person independent pronoun in Grebo from example (2.18). I do not want to conclude too much from one sole example, but my impression is that the zero in Grebo is indeed better interpreted as non-existing. This is also the way it is dealt with in the description of Grebo (Innes, 1966:50,61).

absence of any marking because it is also possible for first or second person subject to be zero, as shown in (2.21 d). It has to be contextually determined who is the subject of sentences like (2.21 c,d); there is no unambiguous linguistic clue (Dixon, 1977:165-194).

(2.21) YIDIJ

- a. *nayu* *ɖuŋga:na*
 1SG,PRON run,PURP
 ‘I had to run’ (Dixon, 1977:527)
- b. *nundu* *gana* *wayga:ɖin*
 2SG,PRON try get up,IMPERF
 ‘You try to wake up’ (Dixon, 1977:516)
- c. *ɖuŋga:n*
 run,PAST
 ‘He ran away’ (Dixon, 1977:528)
- d. *bandi:ldaju* *bama* *ɲabi*
 find,come person many
 ‘I came and found lots of people’ (Dixon, 1977:531)

The crucial argument for the existence of a zero form is that the counterpart (here: speaker and addressee marking) is obligatorily marked in a certain ‘slot’ of the linguistics structure. The emptiness of this slot can then be interpreted as having a meaning.

‘[A zero] is therefore possible only under specific circumstances and in a specific context, namely one that allows, or even favours, the evocation of the (absent) paradigmatic counterpart. A hearer, however, can be expected to notice an absence ‘automatically’ only if the missing counterpart is itself ‘automatic’ – ie very probable – in the particular context. In short, a syntactic zero is possible only when the non-occurring (positive) alternative is so likely that it in fact constitutes the norm.’ (Garcia & van Putte, 1989:369-370)

In case of inflectional marking, this is obviously true. Inflectional paradigms have –by definition– a fixed linguistic ‘slot’ and the non-occurrence of any marking in this slot can be given a linguistic interpretation. In the case of independent pronouns, this is not necessarily the case. Often, independent pronouns are relatively free to occur in various places the sentence, and the absence of a pronoun is not a linguistic sign. As is illustrated by the example from Yidij, the absence of an independent pronoun cannot be interpreted as a reference to third person. Consequently, languages that have optionally marked independent pronouns with a ‘zero’ third person are better analysed as having a ‘non-existing’ third person.

These non-existing ‘zero’ independent third person pronouns are frequent among the languages of the world. A thorough inspection of this phenomenon would have to compare the independent pronouns with the demonstratives of a language, because often the same elements are used, but the homophony is not always noted explicitly by grammarians. As the study of demonstratives falls outside the scope of the present work, I do not have a complete survey of the distribution of this phenomenon in the languages of the world. As a conclusion, I will just have a few random jumps through the languages of the world, and end with some quotes and references to languages

where demonstratives are used for third person reference. Demonstratives for third person reference are found, for example, in Turkish (Lewis, 1967:67-68, 71) and in the Nakh-Dagestanian languages in the Caucasus.³² It is not common in Indo-European, but it does occur in Indo-Aryan:

‘A number of [Indo-Aryan] languages lack 3rd person personal pronouns, using deictics (demonstratives) instead. These are differentiated for a minimum of two categories (distant/proximate) and in some languages for as many as four: eg, Sinhalese *mēka/ōka/arake/ēka* ‘this one, that one (by you), that one (yonder) that one (not in sight, only spoken about)’. They may be differentiated for gender as well.’ (Masica, 1991:251)

An example from New Guinea is found in Yimas, a Sepik Ramu language from Papua New Guinea:

‘True pronouns in Yimas only belong to the first-and second-persons, ie refer to the speaker and addressee. The so-called third-person pronouns refer to participants absent, or not directly interacting in the immediate, ongoing speech act, and belong to the word class of deictics in Yimas.’ (Foley, 1991:111)

In Australia, non-existing third person pronouns are attested regularly:

‘Many languages are like Dyirbal in having just a set of definite determiners – with obligatory specification of distance – in place of anything that could be called ‘third person (singular)’. Sometimes there are also 3 dual and 3 plural pronouns, as in Dyirbal, but a number of languages lack even these.’ (Dixon, 1980:357)

Finally, in a survey of the linguistic structures of the languages of Central America, it is commented that:

‘... formally third person is frequently expressed not by a pronoun but by a demonstrative.’ (Suárez, 1983:81)

Demonstrative third persons are found often. In all those cases, the concept ‘zero’ independent pronoun is at the very least a highly troublesome one. More daringly put: zero independent third person pronouns do not exist.

2.7 Conclusion

Much is possible in human language. There are numerous possibilities to mark singular participants, even when only the three basic categories – ‘speaker’, ‘addressee’ and ‘other’ – are considered. A few general tendencies have been extracted out of the diversity. First, it has been shown that homophony between any two of the three categories is possible, such homophonous paradigms are not numerous among the languages of the world. Examples of almost all logical possibilities exist (only one of nine possibilities has not been attested). Examples of singular homophony only exist in inflectional paradigms, or, in a reversed formulation, independent pronouns do not show homophony. For homophonous inflectional paradigms, it is often argued that it is necessary to add independent pronouns for clarification, as these ‘impoverished’ paradigms have lost some of disambiguating power. As I have shown, this is not necessarily the case. In most cases, languages with singular homophony are still ‘pro-drop’. They do not obligatorily disambiguate the conflation of second and third person. Further, when all three categories are marked differently, then it is highly unusual

³² Eg Hunzib (van den Berg, 1995:60), Lezgian (Haspelmath, 1993:184,190,401-404) and Tsez (Comrie *et al.*, 1999).

to use a zero for the marking of speaker or addressee. More commonly, the ‘other’ category is marked zero. However, in the case of an unmarked ‘other’ it is difficult to argue for the existence of the zero form. Only in the case of inflectional marking, can good arguments be made for the existence of a zero ‘other’. In the case of independent pronouns, a zero ‘other’ is probably better interpreted as the non-existence of a third person pronoun.

Now the groundwork has been done, it is time to proceed and expand the investigation of person marking on these foundations. The three singular categories that have been the subject of this chapter (‘speaker’, ‘addressee’ and ‘other’) will now be used to approach the marking of what has traditionally been called ‘plural’ pronominals.

Chapter 3

Group marking

Redefining plurality in the pronominal domain

3.1 Introduction

After singular comes non-singular. In this chapter, the first – cautious – steps are taken into the domain of non-singular pronominal reference. The diversity of non-singular pronominal marking is so extensive that all the following chapters will deal with some aspect of this theme. In this chapter, the basic framework for the analysis of non-singular marking will be developed. This framework is built on the basis of the singular categories from the previous chapter. The major part of this chapter will be devoted to the question how to deal with the marking of groups of participants. Morphemes that refer to such a group of participants are normally referred to as ‘plural’ pronominal elements. The main objective of this chapter is to redefine the notion of ‘plurality’ in the pronominal domain. The notion of ‘group’ marking is proposed to replace ‘plurality’. This perspective is not completely different from the traditional one. The difference between ‘plural’ and ‘group’ marking only amounts to a small change, but the new perspective allows for a cross-linguistically more sensible typology. A few consequences of the new definition are discussed in this chapter, but the full typology on the basis of the new perspective is postponed to the next chapter.

This chapter starts with a discussion of the definition of the phenomena to be reviewed. In general, a cross-linguistic comparison only makes sense when the objects of comparison are comparable. This implies that the definition of the boundaries is one of the most important aspects of a comparison. Once the boundaries are defined properly, the actual comparison follows almost automatically. These definitional questions will be discussed in section 3.2. Next, the terminology that is needed to talk about non-singular marking will be discussed in section 3.3. In this section, the known problems with the traditional notion ‘plural’ are summarised. The notion ‘group’ is proposed to improve on the deficiencies of ‘plural’. Equipped with a proper definition and terminology of ‘group’ marking, the possible categories within these definitional boundaries are reviewed in section 3.4. A complete typology on the basis of the possible categories turns out to be rather laborious. Therefore, only part of the complete typology will be taken up in this chapter. In section 3.5, the typology of the ‘first person complex’ is discussed. The first person complex consists of the categories that are subsumed under the meaning of the English pronoun ‘we’. The typology of the other group-categories is postponed till the following chapter. As for this chap-

ter, the different attested marking patterns of the ‘first person complex’ in the world’s languages are discussed in section 3.6. Some generalisations are drawn on the basis of this partial typology in section 3.7. Finally, the content of this chapter is summarised in section 3.8.

3.2 Definition

The previous chapter discussed the marking of pronominal elements with singular reference. The subject of the present chapter is the marking of pronominal elements with non-singular reference. However, this definition has to be more strictly specified to be able to restrict the diverse manifestation of pronominal marking in the world’s languages to a workable format. The basic definitional property for a pronominal morpheme to be included in the present discussion is that it has to refer to more than one person or object. Four extra criteria will be added onto this basic definition. Summarised shortly, these four criteria are the following. First, the multiple persons or objects have to be regarded as a group. Second, the morphemes have to be unmarked to the specific amount of elements. Third, the morphemes should not include the regular reference for singular persons or objects. Finally, paradigms with secondary distinctions, like gender or honorifics, will be disregarded. These four criteria will now be discussed in turn.

The first extra criterion is that the multiple persons or objects are to be interpreted as a group. A sentence like ‘the students lifted the heavy table’ refers to a situation where the students form a group and act together (the subject of the sentence has collective reference). However, it is not necessary to act together to form a group. A sentence like ‘the students passed the exam’ normally means that all students individually passed the exam (it has a distributive reference). Still, ‘the students’ in the second sentence are defined as one group in the present study. The only condition to form a group is that the persons or objects are in the same predicative role. This condition is important because there are languages that have pronominal elements that refer to more than one person, but these persons are not in the same predicative role. An example is presented in (3.1), from the Yuman language Mojave. The pronominal prefix n^v -... refers to two persons. In one possible meaning, the two persons are first person subject and second person object (glossed as ‘1→2’). In another possible meaning of the prefix n^v -..., the two persons are third person subject and first person object (glossed as ‘3→1’). No other meaning is possible for this prefix. Concluding, the pronominal prefix n^v -... refers to more than one person or object, so it fits in with the basic property of the definition. However, the persons or objects referred to by n^v -... are not in the same predicative role, so this prefix does not hold to the first extra criterion of the definition.

(3.1) MOJAVE

n^v -masde:k

1→2/3→1 – afraid

‘I’m afraid of you’, ‘He’s afraid of me’

(Munro, 1976: 12)

The next two extra criteria are concerned with the specification of the number of participants. The basic definition states that reference should be made to more than one person or object. However, the specific number of persons or objects should not be specified. Special pronominal elements for reference to a specific number of persons, like dual, trial or paucal, are excluded here. These elements will be taken up again in part Three. In this chapter, only those elements are included that are unmarked for the number of participants. Also, elements that can refer to singular participants are excluded. This is an important restriction. Often, there are no grammaticalised plural forms in a pronominal paradigm. The singular forms are used for the plural as well. For example in Mojave, the pronominal prefixes distinguish first, second and third person singular. However, there are no special prefixes that mark a group of more than one person. The same singular prefixes are used for plural marking. As shown in (3.2), there is often no number marking at all in a Mojave sentence. The subject can be either singular ‘I’ or plural ‘we’. The meaning has to be recovered from the discourse. It is possible in the Yuman languages to specify plurality, but it is never marked by a pronominal element.¹

(3.2) MOJAVE

ʔ-aherk

1-put in jail

‘I put him in jail’, ‘we put him in jail’

(Munro, 1976: 15)

Elements like the prefix *ʔ-...* of Mojave, that both mark for singular participants and for groups of participants, are not included in this chapter.² Only *specialised* pronominal elements for group marking are included. A consequence of this strict definition is that there are paradigms, like the pronominal prefixes from Mojave, that do not have any (specialised) morphemes for group marking.

There is a last criterion to restrict the phenomena to be discussed in this chapter. This restriction was already introduced in the previous chapter (see section 2.2); I will only briefly repeat it here. All pronominal paradigms that include sub-categories, like gender or honorifics, are excluded. The status of these sub-categories within pronominal paradigms is a problem that deserves a study in its own right. To achieve a clear typology of personal pronominal marking, I have chosen not to include these complicating factors in this survey. For now, and for the following chapters, I disregard all pronominal paradigms that make secondary distinctions in the referential categorisa-

¹ The marking of plurality in the Yuman languages is quite complex, using affixes, stem changes or even completely different verbal stems for singular and plural subjects. All in all, there are about 25 different ways to mark plurality (eg on Mojave see Munro, 1976:224-232), the distribution of which is still a puzzle to the specialists:

‘Some efforts have been made in previous descriptions of Yuman languages to class certain plural markers as indicating collective as opposed to distributive plural ... or to refer to semantic similarities among the verbs which take a given plural marker. ... However, I do not feel that any of these typologies adequately explains the puzzling distribution of plural markers.’
(Munro, 1976:227)

² This criterion is the counterpart to one of the criteria used for the definition of singular marking in the previous chapter. In section 2.2, pronominal elements that cross the singular/plural boundary were considered to be singular. In this chapter, the same criterion is used to exclude these elements in the discussion of plurals.

tion of pronominal elements. In a few incidental cases, I deviate from this restriction to discuss a very special case of paradigmatic structure. These cases, all with a gender distinction in the third person singular, will be explicitly noted in the discussion of the examples.³

3.3 Terminology

Now that the definition is clarified, it is time to take a look at the terminology. Over the last century, there has been an ongoing, but rather covert, discussion about a suitable meta-language to approach the marking of person in human language. Many deficits of the traditional approach have been noted. However, there is still no widely accepted replacement for the traditional term ‘plural’, which turns out, on closer inspection, to be a rather awkward concept for the description of non-singular pronominal categories. I will make an attempt for the long overdue reformulation by replacing ‘plural’ with ‘group’.⁴ The reasons for this reformulation will be laid out in this section. All that follows in the present work can be read as a prolonged argument for the new terminology.

Pronominal elements that refer to a group of persons or objects, as defined in the previous section, are normally called ‘plural’. However, there is a problem with the term ‘plural’. The notion of ‘plural’ within the domain of pronominal marking is rather different from the standard notion of ‘plural’. Normally, a singular morpheme, like the English word ‘chair’, refers to a single object that falls into the class of chairs. A plural, like ‘chairs’, refers to a group of objects, each of which belongs to the class of chairs. Transferred to the pronominal domain, this would mean that a ‘first person singular’ refers to a person that belongs to the class of speakers. No problem up to here. However, the consequent next step would be that a ‘first person plural’ refers to a group of persons that all individually belong to the class of speakers. In other words, a ‘first person plural’ is literally a ‘group of speakers’. The English pronoun ‘we’ would, in this analysis, mean something like ‘the group of persons currently speaking in unison’. This is clearly not what ‘we’ prototypically means. ‘We’ normally refers to a group of people, only one of whom is currently speaking (Jespersen, 1924:192; Benveniste, 1966:233; Lyons, 1968:277; Moravcsik, 1978:354, fn. 12). The most common meaning of ‘we’ strongly resembles the meaning of a nominal case marker that is known as the ASSOCIATIVE. An associative case marker is, for example, found in Hungarian. Hungarian has a nominal plural suffix *...-ok*, as in (3.3a). This phrase can, for example, be used to refer to a group of people that all have the name John. There is a different associative suffix *...-ék*, as in (3.3b). With this suffix, the phrase means something like ‘John and his group’.

³ This restriction could in principle lead to a skewed picture of the distribution of the different types of pronominal paradigms. However, my impression is that this is not the case. When I came across a pronominal paradigm with gender or honorific distinctions, I have always checked whether the paradigm did show an unusual structure, not attested in any other paradigm. It never did. Also, my impression is that no type is grossly underrepresented in the current set of examples because of the omission of paradigms with special gender or honorific categories.

⁴ Howe (1996: 9) also uses the term ‘group’ instead of the term ‘plural’.

(3.3) HUNGARIAN

a. *János-ok*

NAME-PLUR

‘More than one John’

b. *János-ék*

NAME-ASSOC

‘John and associates’, ‘John and his group’ (Corbett & Mithun, 1996:5)

The prototypical meaning of ‘we’ is associative. ‘We’ means something like ‘I and my associates’, in which the associates could be either addressees or others. A real ‘plural I’ is possible, but rather odd. For the ‘second person plural’, the situation is different. Here, both the plural reference ‘you all, present audience’ and the associative reference ‘you and your associates’ make sense. The prototypical meaning of the ‘second person plural’ is probably associative (see section 3.4 for an extensive discussion). For the ‘third person plural’, both the plural and the associative meaning are completely normal, but the prototypical meaning is probably plural (Moravcsik, 1994). The meaning of the so-called ‘plural’ pronominal elements contain both plural and associative notions. This makes the term ‘plural’ in the domain of pronominal marking semantically awkward and somewhat confusing. It is better to avoid this term.

This theoretical argument can be supplemented with an empirical observation. It is rather uncommon to find pronominal paradigms that use nominal strategies to mark plurality. Normally, the ‘plural’ in the pronominal domain is marked differently from the plural of regular nouns in the language. Benveniste (1966) already made this observation.

‘... l’identité des formes pronominales au singulier et au pluriel demeure l’exception.

Dans la grande majorité des langues, le pluriel pronominal ne coïncide pas avec le pluriel nominal, du moins tel qu’on le représente ordinairement.’ (Benveniste, 1966:233)

Examples of nominal plural marking in pronouns exist, but they are generally hard to find. Note that it is rather common to find pronouns with transparent plural forms; plurals that are built by regular derivation from the singular forms. However, the PRONOMINAL marking for plurality is in most cases (synchronically) different from the NOMINAL marking for plurality.⁵ In those few cases where a nominal marker is found in the pronominal paradigm, it is either found in part of the paradigm only, or it is functionally empty (ie the singular and plural forms are still different if the plural marker would be removed), or it is only optionally used. It is really uncommon for a pronominal paradigm to use a nominal plural marker as the *only* identification of plurality. An example of such a rarity is found in Trumai, to be discussed below. After the case from Trumai, I discuss two cases (from Korean and Canela-Kraho) that come near, but even in these cases the pronominal plural marking is not completely regular.

⁵ Regular pronominal marking for plurality is commonly related to nouns like ‘fellow’ or ‘people’. The marking ‘I-fellow’ or ‘I-people’ should probably be read as marking an associative case: ‘I and my fellows’ or ‘I and my people’ (cf section 4.4.2).

An instance of a complete overlap between pronominal and nominal number marking is attested in Trumai, a genetically isolated language from Brazil (Monod-Becquelin, 1975: 175). The person markers *ha* and *hi* (meaning ‘I’ and ‘you’ respectively) are combined with the plural postposition *uan* to form ‘we’ and ‘you-plural’. The same pluraliser is also used with nouns. From the accompanying texts (Monod-Becquelin, 1975:226-247), it appears that this pluraliser is used obligatory with pronouns and human nouns, but it is not consistently used with non-human nouns.

‘Il est clair que les deux personnels <ha> et <hi> se combinent avec le pluralisateur = <ha uan> = 1^{re} personne + pluriel = nous; comme <kiki>, <un homme>, <kiki uan> = <les hommes>. (Monod-Becquelin, 1975:175)

In Korean, the plural marking is used obligatorily in polite levels of address in all but the first person plural:

‘In non-deferential levels of address, there is a separate stem for 1st plural; However, the other persons use the same stem [as the singular, MC], but are obligatorily marked by a number suffix.’ (Sohn, 1994:284)

For example, in the ‘level II’ pronouns in Korean (used when the addressee of third person referent is generally equal or lower than the speaker, yet he or she is an adult) the second person singular pronoun is *tangsin* and the second person plural pronoun is *tangsin-tul* (Sohn, 1994:287). The suffix *...-tul* is the regular number marker for nouns, used obligatorily when the noun is definite (Sohn, 1994:268-269). However, the first person plural is *wuli*; rather different from the first person singular *ha*. *Wuli* is only optionally extended with the nominal suffix *...-tul*. Thus, it can be concluded that there is some nominal plural marking in pronouns, but only in a restricted part of the paradigm.⁶

Another example where pronominal elements use the same plural marking as regular nouns is found in Canela-Kraho, a Gé language from Brazil. The plural particle *me* is found with pronominal arguments, as in (3.4b), as well as with nominal arguments, as in (3.4c). However, this particle is not obligatorily used in the language and even when it is used, the context has to make clear whether it is the subject or the object that is meant to be plural.

‘Number is sometimes expressed by the particle ‘me’, ‘plural’, usually where the referent is human and more specifically, Indian. ... The context alone indicates whether the subject or the object is being pluralized.’ (Popjes & Popjes, 1986:185)

(3.4) CANELA-KRAHO

a.	<i>Capi</i>	<i>te</i>	<i>i-pupun</i>	
	NAME	PAST	1-see	
	‘Capi saw me’			(Popjes & Popjes, 1986: 175)

b.	<i>Capi</i>	<i>te</i>	<i>me</i>	<i>i-pupun</i>	
	NAME	PAST	PLUR	1-see	
	‘Capi saw us (exclusive)’				(Popjes & Popjes, 1986: 175)

⁶ The independent pronouns from the Siouan language Lakhota are comparable in structure to these Korean pronouns. The Lakhota pronouns are discussed in section 3.6.6, example (3.18). The plural suffix *...-pi* that is found in the pronouns is also used occasionally to pluralise nouns, although it is not suffixed on the noun, but on the verb (Van Valin, 1977:60-62).

c.	<i>hūmre</i>	<i>te</i>	<i>me</i>	<i>rop</i>	<i>cahhyr</i>
	man	PAST	PLUR	dog	beat

‘The men beat the dog’

(Popjes & Popjes, 1986: 186)

To sum up, the marking of non-singular pronominal elements is neither semantically nor morphologically similar to nominal plural marking. A much better approach for the analysis of pronominal paradigms is to talk about ‘group marking’ instead of ‘plural marking’. This may seem like a terminological quibble, but this slight change in perspective leads to a rather different typology; one that is better suited to tackle the cross-linguistic variation of ‘plural’ pronominal marking. Basically, the change in perspective from plural marking to group marking is a change in emphasis from NUMBER to KIND. In different words, a change will be proposed from a QUANTITATIVE to a QUALITATIVE criterion.

The traditional terminology of ‘plural’ marking highlights the number of participants: there are singular (one) and plural (more than one) pronouns. The difference between the two is one of quantity. From this perspective, there are three basic singular participants (speaker, addressee and other) and, likewise, there are also three basic plural participants (speaker-plural, addressee-plural and other-plural) – in total six kinds of pronominal elements. This traditional classification is not only semantically awkward, as set out above, it also gets tangled up when it incorporates the difference between an inclusive and an exclusive first person plural. The difference between an inclusive (a first person plural that includes the addressee) and an exclusive (a first person plural that excludes the addressee) is found regularly among the world’s languages. However, as this distinction has no place in the basic six-way typology, a new dimension is introduced to take care of this opposition. However, the inclusive-exclusive opposition is not at all independent of the six-way typology. The inclusive-exclusive opposition is only found in the first person plural. There is no inclusive-exclusive distinction in the second, nor in the third person. A complete new dimension to account for a single extra form is rather a waste of apparatus.

The perspective that will be taken here is a different one. It does not emphasise the number of participants, it emphasises the kind of participants in a group. In this view, there are groups of participants, as opposed to singular participants. These groups are analysed according to the kind of participants of which it is composed. A group, for example, can consist of a speaker and the addressee, or of the addressee and some other participant, et cetera. The number is not important, only the kind of participants involved is important. The group-perspective is a qualitative approach to non-singular marking. The group-perspective on pronominal elements, which will be developed in the next section, is semantically less awkward, and it will incorporate the inclusive-exclusive distinction automatically. A consequence of this approach is that a group could also consist of multiple speakers, or of multiple addressees. Both these ‘real plural’ meanings are part of the logically possible groups of participants. However, these possibilities will be thrown out of the typology, on empirical grounds, in the next section. These ‘real plurals’ do not occur as grammaticalised elements in the languages of the world.

3.4 Towards a typology of groups

Based on the threefold division of singular participants (1: speaker, 2: addressee, 3: other), there are seven logical possibilities to form groups. The seven possibilities are shown in Table 3.1. The numbers in the table refer to the kind of participants, the amount of each kind is not important.⁷

1+1	‘we’, mass speaking
1+2	‘we’, including addressee, excluding other
1+3	‘we’, including other, excluding addressee
1+2+3	‘we’, complete
2+2	‘you-all’, only present audience
2+3	‘you-all’, addressee(s) and others
3+3	‘they’

Table 3.1: Different groups of participants⁸

Of the seven possible groups, there are four that include the speaker; 1+1, 1+2, 1+3 and 1+2+3. These are all subsumed under the meaning of English ‘we’ (Quirk *et al.*, 1985:340). Franz Boas was the first to distinguish these four possible meanings of the English pronoun ‘we’. He also concludes that a ‘true first person plural’ (ie 1+1) is impossible.

‘When, therefore, we speak of a first person plural, we mean logically either self and person addressed [1+2], or self and person or persons spoken of [1+3], or finally, self, person or persons addressed, and person or persons spoken of [1+2+3]. *A true first person plural [1+1] is impossible, because there can never be more than one self.*’ (Boas, 1911b:39, number-notation and italics added, MC)

However, interpreted slightly different, the ‘true first person plural’ 1+1 exists as a conceptual category. Boas interpreted the ‘first person’ as reference to the SELF, and of course, only one ‘self’ can speak at a time.⁹ When ‘first person’ is interpreted as reference to the SPEAKER, as is done here, then a ‘true first person plural’ (1+1) can be interpreted as a group of speakers, speaking together in unison. Mühlhäusler & Harré

⁷ Cf Sokolovskaya (1980:90), Plank (1985:130-152), Henderson (1985), Greenberg (1988:13-15) and Noyer (1992:147-148) for comparable approaches, but using slightly different notations.

⁸ It is an old practice to abbreviate inclusive and exclusive with the number 1 & 2 and 1 & 3 respectively. Jespersen (1924:213) remarks that the notation ‘1/2’ and ‘1/3’ has been used in ‘some works dealing with Amerindian languages’. I have been unable to track down an instance of this notation. The first complete analysis of plural pronouns as combinations of singular participants is probably by Hollenbach (1970), although she does not use the ‘+’ sign. An early usage of the ‘+’ sign for inclusive and exclusive first person plurals can be found in Lounsbury (1953). The ‘+’ sign is for the first time consistently used by Zwicky (1977). The notation with a ‘+’ has become en vogue to analyse plural pronominal elements in the descriptive literature from the eighties onward (eg Roberts, 1987:210-211). It is unclear to me what has given it this sudden boost. There is some variation in the current literature, as sometimes ‘1+1’ is found for the first person plural exclusive, where I use ‘1+3’, eg various articles in Wiesemann (1986a) and also in Hale (1973:315).

⁹ The ‘self’ is a much debated concept in philosophy and ethnology. For a summary of the relevant aspects of this debate for linguistics, see Mühlhäusler and Harré (1990:104-122) and Foley (1997:261 ff.). For the analysis of linguistic structure, the concept of the ‘self’ is a theoretical burden, rather than a blessing. The concept ‘speaker’ is more appropriate for the present purpose. Alternatively, for the analysis of language use, the concept ‘self’ is indispensable.

(1990:201-205) note that the use of ‘we’ in the meaning of a ‘group of persons speaking in unison’ is found in the phenomenon of MASS SPEAKING:

‘... football chanting, ritual mass speaking, as in a church service, the mass speaking of children at play and finally the reactions of a concert audience. There are many other forms of mass speaking, such as what occurs at political rallies. ... In order for there to be mass speakings at all, the members of the speaking group or groups must achieve a high degree of co-ordination of their actions.’ (Mühlhäusler & Harré, 1990:201-202)

An example is *we are the champions*, sung by a soccer audience after their team has won, or *we want more* shouted by a concert audience. This is a rather marked usage of the English pronoun ‘we’, but it can be used that way. Still, this does not make 1+1 a linguistic category. For it to be a linguistic category, the notion of ‘mass speaking’ should exist as a grammaticalised category at least in some languages in the world. However, as far as I know, there is no language in the world that distinguishes a separate morpheme for ‘mass speaking’. It seems that Boas was on the right track when he proclaimed that 1+1 was an impossible category. His rationale does not hold, but his intuition was right. 1+1 is semantically a feasible category, but it is not linguistically salient. For the rest of the discussion here, I will disregard this category altogether.¹⁰

Just as 1+1, the category 2+2 is not widely recognised as a viable linguistic category. It is a viable category at least in principle. The category 2+2 is a semantically feasible category. When addressing a group a people, there is in principle a difference whether ‘you (more than one person)’ addresses only the present audience and no one else, or whether other people, not directly addressed, can be included. The reference only to the *present audience* is indicated here by 2+2; the general group address is indicated by 2+3. Lyons describes this difference as follows:

‘A distinction might also be made between an ‘inclusive’ and ‘exclusive’ use of the ‘second person plural’ (in a slightly different sense of ‘inclusive’ vs. ‘exclusive’). The English pronoun *you* may of course be either singular or plural As a plural form, it may be either ‘inclusive’ (referring only to the hearers present – in which case is the plural of the singular *you*, in the same sense as *cows* is the plural of *cow*) or ‘exclusive’ (referring to some other person, or persons, in addition to the hearer, or hearers).’ (Lyons, 1968:277)

A lot of the situations, where in English ‘you (more than one person)’ is used, are instances of 2+2. One could think of a situation like school class address. A teacher asking *should I tell you more about this?* in class, is referring only to the present audience (2+2), not to all his students, some of which might be absent. However great the semantic plausibility, the category 2+2 is not found grammaticalised in the languages of the world (Moravcsik, 1978:356; Greenberg, 1988:14; McGregor, 1989:450, fn. 8). I have found one source where a morphological difference between 2+2 and 2+3 is claimed. Hewitt (1979) notes that in the Caucasian language Abkhaz, a difference is made between an exclusive and an inclusive ‘second person plural’. The exclusive form *šʷart* indicates ‘you excluding them’, which ends up identical in meaning to what I have called 2+2. The inclusive form *šʷa(rà)* indicates ‘you including them’; what I have called 2+3.¹¹ This opposition is not very strong within the language: the

¹⁰ Cf Zwicky (1977:731, fn.1) for the same reasoning.

¹¹ Hewitt’s analysis of an inclusive-exclusive opposition in the second person plural probably goes back to Š.K. Aristeva et al. (1968) *Grammatika abxazskogo jazyka; fonetika i morfoložija* [Grammar
[continued on next page]

exclusive forms are hardly used; the ‘inclusive forms may and usually do appear where one would expect the exclusives’ (Hewitt, 1979: 156-157). Also, this opposition remains a special characteristic of Abkhaz; it is not found in any of the other Northwest Caucasian languages.¹² On the whole, it seems rather implausible, if not impossible, for a language to have a grammaticalised category 2+2.¹³ I will disregard this category from now on.

The categories 1+1 and 2+2 are thus possible categories in principle, but they are not attested as linguistic categories. These semantic categories do not grammaticalise in human language. This can be explained by noting that the conversational settings in which the semantic categories 1+1 and 2+2 are attested are extremely marked. Speech in unison must be carefully planned, and eye contact with more than one person is only achieved in a specific situation, like class-address (McGregor, 1989: 440). Both in the case of multiple speakers (mass speaking) and of multiple addressees (present audience), the group has to act as if it were a single entity. The multiple speakers have to ‘speak as one’, and the multiple addressees have to ‘be conceived as one’. For a linguistic theory of pronominal marking, it is not necessary to consider combinations like 1+1 or 2+2 as separate categories. In this way, a remarkable reduction of the meta-languages as proposed by Plank (1985: 130-152) and Greenberg (1988: 13-15) can be achieved. They both distinguish all cases with multiple occurrence of the number 1, 2 and 3 as feasible linguistic categories. However, as 1+1 is unnecessary, so is 1+1+3, and 1+1+1+3, et cetera. By the exclusion of 2+2, the same argument can be made for the exclusion of 2+2+3, 2+2+2+3, et cetera. Also, combinations like 1+1+2+3 or 1+2+2+3 are unnecessary, for the same reason. The fact that the categories 1+1 and 2+2 are not necessary for a linguistic theory of pronominal marking substantiates the group-approach as outlined in Table 3.1.¹⁴

[continued from previous page]

of the Abkhaz language; phonetics and morphology] Suxumi, Alašara. On page 35, it says that *hara* is inclusive first person plural and *hart* is exclusive first person plural. Then it is commented that ‘The second person plural is formed analogously: *shvara*, ‘you-plural’ and *shvart*, ‘you-plural’.’ Thanks to R. Smeets for finding the original reference. He also explains that the effect of the suffix *...-t* is probably comparable to the function of the French postposition *autres* in *nous autres* en *vous autres*. The same suffix *...-t* is found in the Abkhaz demonstrative elements, but then in the function of a pluraliser.

¹² Abkhaz is part of the Northwest Caucasian family. The difference between an inclusive and an exclusive second person plural is not noted for any of the other Northwest Caucasian languages (cf the descriptions in Hewitt, 1989: 101, 176-177, 298, 378).

¹³ Comrie (1980b: 837) mentions Southeast Ambrym as an example of a language with the same peculiarity as Abkhaz. He does not give any reference to substantiate this claim. Plank (1985: 147) takes over the claim on face value. The claim probably goes back to Parker (1970: ix, 43) who lists two different forms for ‘you-plural’ in his dictionary of Southeast Ambrym: *xami*, which is glossed as ‘you plural inclusive’ and *xamim*, which is glossed as ‘you plural exclusive’. However, I have not been able to find any description of this phenomenon in the literature on other variants of Ambrym (cf Paton, 1971: 13-18, 119-124).

¹⁴ In a reply to a comparable argument by McGregor (1989), Greenberg (1989) raises the important issue of the function of a typological classification. He argues that a classification is not only a description of attested possibilities, but that it should also point towards distinctions which might exist, but have not been noted until now because the linguistic meta-language did not allow them to be described.

‘It is a fact of ordinary experience, that a single speaker often addresses several hearers simultaneously both on informal occasions and in public address. I cannot, in fact, produce a single example of a language which does make a clear distinction between (2+2...) and (2+3...), but [it]

[continued on next page]

Five categories remain after these two have been dismissed. Three of the remaining categories include the speaker. These are 1+2, 1+3 and 1+2+3. These categories all exist as grammaticalised categories in the world's languages. However, in English, they are not distinguished morphologically. All three have to be translated into English by the pronoun 'we'. They will be quickly introduced here. The categories 1+2 and 1+2+3 include reference to the addressee. They are both known as an 'inclusive first person plural'. In English, there is no specialised pronominal element for these referential values. To express its content, a paraphrase like 'we, you also' has to be used. Some pronominal paradigms in the world's languages distinguish between the two categories 1+2+3 and 1+2. In such paradigms, the category 1+2+3 refers to a group that includes the speaker, the addressee and at least one other. This is known in the literature as an 'augmented inclusive'. On the other hand, the category 1+2 is known in the literature as the 'minimal inclusive'. In this case, only the present audience is included, prototypically only one singular addressee. More discussion of these categories will be presented in section 3.6. The category 1+3 refers to a group excluding the addressee. This is generally known as the 'exclusive first person plural'. In English, it has to be paraphrased as something like 'we, I and the ones I was talking about, not including you'. The other two categories are well-known and do not need of explanation. The categories 2+3 and 3+3 are equivalent to the 'second person plural' and 'third person plural', respectively. As far as these categories are concerned, the 'group' perspective is not much different from the 'plural' perspective.

3.5 A partial typology: the first person complex

The five different group categories – 1+2+3, 1+2, 1+3, 2+3 and 3+3 – can be marked linguistically in numerous different ways. At one side of the spectrum, they can all five be marked with a different morpheme. That would amount to a maximally differentiated paradigm. At the other extreme, none of the categories could have a specialised morpheme. All group reference could be marked using only the singular pronominal elements and no other morphemes. That would amount to a minimally differentiated paradigm. In between these two extremes there are all kinds of different ways to group the categories together. Counting all the possible combinations, there turn out to be 203 theoretically possible patterns.¹⁵ To find a way through the bewildering

[continued from previous page]

is obviously an empirical possibility to whose occurrence we should be alert. (Greenberg, 1989:454, italics added, MC)

I agree that we should be alert to hypothetical other distinctions, but in this case, it seems to me rather a waste of apparatus to include all the encumbering multiple combinations of 1 and 2 in the meta-language only because of some non-attested possibilities.

Besides this theoretical argument, Greenberg presents a few languages to argue for the necessity of multiple addressees in the meta-language. However, these examples do not convince me. The cases of Palaung and Sierra Popoluca are interesting, but they do not argue for multiple addressees (see section 8.5.5 and 4.5.2, respectively, for an extensive discussion of these cases). A third case presented by Greenberg concerns the Philippine pronouns. The Pampangan pronoun *tamu* is probably built from the parts *ta* (1+2) and *mu* (2), but its meaning is clearly 1+2+3, not 1+2+2 as Greenberg wants to have it (Gonzales, 1981:172). Also, the equivalent pronoun *tayo* in the related language Ilocano is built from the parts *ta* (1+2) and *yo* (2+3), see also section 3.6.5.

¹⁵ The total of 203 possible combinations is calculated by taking the sixth so-called 'Bell number'. A Bell number B_n represents the total number of possible dissections of a set of n different elements. In

[continued on next page]

amount of 203 feasible patterns, I will start off with a subset of the five categories. The typology will be slowly built up from there. The point of departure will be the ‘first person complex’; the set of categories that include the speaker. There are three categories that include the speaker, 1+2, 1+3 and 1+2+3.¹⁶ With only three categories, and the possibility of non-existence, there are 15 possible patterns.¹⁷ In this section, I will show which of these 15 patterns are attested among the world’s languages. In section 3.6, I will give a description of the attested patterns. I will draw some generalisations on the basis of the attested patterns in section 3.7. The distribution of the other two categories – 2+3 and 3+3 – relative to the patterns of the first person complex is a long story. This will be taken up in the next chapter. All patterns that are theoretically possible with the three categories from the first person complex are shown in Table 3.2. A capital letter in this table indicates that there is a specialised morpheme for this category – or combination of categories. A dash indicates that this category is linguistically coded for by a morpheme that also marks a singular category. In case of a dash, there is no specialised morpheme for the referential category. The possibilities are ordered according to the number of specialised morphemes that are present. The first pattern distinguishes three different morphemes for the three different categories. Next follow the six possible patterns with two specialised morphemes. Then the seven possibilities with only one specialised non-singular morpheme are shown. Finally the only possible pattern with no specialised morpheme for either of the three non-singular categories closes the list of possible patterns. In total, this makes 15 different patterns. I have found ten of these 15 patterns in the world’s linguistic variation – they are marked with an arrow in Table 3.2.

1+2	A	A	A	A	A	–	A	–	A	–	–
1+2+3	B	B	A	B	–	A	A	–	A	–	–
1+3	C	A	B	–	B	B	–	A	A	–	–
	↑	↑	↑	↑	↑		↑	↑	↑	↑	↑

Table 3.2: Possible patterns of first person complex

[continued from previous page]

this case there are five different categories (1+2, 1+3, 1+2+3, 2+3 and 3+3) and the possibility of non-existence. The number of possible dissections of these elements is $B_6 = 203$.

¹⁶ It is also possible to start with the ‘second person complex’. This would be the set of all categories that include the second person, 1+2, 1+2+3 and 2+3. Or indeed, it is possible to start from the ‘third person complex’, which consists of the categories 1+3, 1+2+3, 2+3 and 3+3. *A priori*, there is no reason to prefer one of these perspectives over the others. The numeral ‘1’ is only shorthand for ‘speaker’, and there is no reason to start with the speaker for the construction of the typology. The order of the numbers in the notation does not have an inherent meaning. The individual numbers in ‘1+2+3’ are just conveniently placed in numerical order. The groups ‘2+3+1’ or ‘3+2+1’ are identical to ‘1+2+3’. To see what it would bring, I also analysed the data from the viewpoint of the ‘second person complex’, but this didn’t bring any new interesting insights. There is a small group of pronominal paradigms where the discussion would benefit from a ‘second person complex’ perspective (see especially the discussion of ‘vertical’ homophony in paradigms with an inclusive/exclusive opposition in section 5.3.3). However, some beautiful generalisations would be missed from that perspective. I did not consider the ‘third person complex’ perspective seriously. This perspective would not make the discussion much easier, as there are four categories in this complex that can be combined into 52 different patterns. This is too much to take up at once, and still retain a clearly structured discussion.

¹⁷ With three different categories (1+2, 1+3 and 1+2+3) and the possibility of non-existence, the total number of dissections is $B_4 = 15$ (see also footnote 15).

The ten attested patterns are repeated in Table 3.3. They are separated in two groups. The first five are common among the world's languages. They are designated for ease of reference as (Pa) through (Pe). The last five patterns – (Pf) through (Pj) – are rare. The distinction between common and rare is rather clear-cut. There are dozens of examples of the common patterns, but only one, or maximally two, examples of the rare patterns. This indicates that the five 'common' patterns are indeed a rather different sort. They are the preferred structures for pronominal paradigms.

	(Pa)	(Pb)	(Pc)	(Pd)	(Pe)	(Pf)	(Pg)	(Ph)	(Pi)	(Pj)
1+2			A	A	A	A	A	A	–	–
1+2+3	A	–	A	A	B	B	B	A		–
1+3			–	B	C	B	A	–	A	A

Table 3.3: Attested patterns of the first person complex

3.6 Different kinds of 'we'

Ten out of the fifteen possible types for the first person complex are attested. The pattern of the first person complex of these types will be exemplified in this section. The first five types are common. The common patterns ask for more attention than can be dealt with in this chapter. Although the amount of logically possible alternatives is reduced enormously by only looking at these five patterns of the first person complex, these five common patterns still show many different combinations with the other two categories (2+3 and 3+3). The full discussion of the geographical distribution and the structural variability of these common types is postponed till the next chapter. The (extensive) list of examples of the common patterns, as presented in the next chapter, should be a reasonable argument for the commonness of these patterns. In this chapter, only a rough outline of the structure of the common patterns will be presented. The other five types are found only in incidental cases. All paradigms attested of these rare types will be completely described in this section.

3.6.1 Type (Pa): 'unified-we'

The structure of type (Pa) is well known. This is the structure that is found in English independent pronouns, with a single undifferentiated form of 'we'. The meaning of a morpheme like the English 'we' is best characterised as a group of more than one person, including the present speaker, but unmarked as to the other persons in the group. It unifies the referential categories 1+2, 1+2 and 1+2+3 in one morpheme. I will use the name 'unified-we' for this type of paradigm. The complete variability of this type is discussed in section 4.3. Two paradigms of the 'unified-we' type will be given some special attention here, because they are classified differently in the literature. These two paradigms, from Assiniboine and Nyulnyul, are presented as special paradigmatic structures by Greenberg (1988:9) and McGregor (1996:40-41). However, following the definitions as used in the present work, these paradigms are cases of the well-known 'unified-we' type.

Greenberg (1988:9) presented the active prefixes from the Siouan language Assiniboine as an example of a special paradigmatic structure, which he, accordingly, called the ‘Assiniboine-type’ paradigm. There are indeed paradigms of this Assiniboine-type attested among the world’s languages (they will be discussed here as type (Pf) in section 3.6.6), but Assiniboine itself is not among them. The problem is that the special opposition Greenberg reported for Assiniboine is not marked inside the pronominal paradigm itself. The special characteristic of Assiniboine is that there is a difference between the category 1+2 (speaker and addressee, no others) and all other categories in the first person complex. There is thus a difference between ‘we: I and you’ and all other forms of ‘we’. These other forms of ‘we’ consists of both inclusive and exclusive reference. This is indeed a very special division of the first person complex. Two example sentences from Assiniboine are presented in (3.5). Sentence (3.5a) shows an example with 1+2 reference; sentence (3.5b) shows an example with all ‘we’ reference except for the 1+2 reference.

(3.5) ASSINIBOINE

- a. ***yk-áksa***
1PLUR-chop off
 ‘I and you chop off something’ (Levin, 1964:31)
- b. ***yk-ákipta-pi***
1PLUR-argue-PLUR
 ‘We argued with him’ (Levin, 1964:32,93)

However, the crucial difference between the two sentences is not found in the form of the pronominal prefix. The prefix ‘*yk-...*’ is in both cases identical. The difference between the two kinds of reference is marked by the presence or absence of the plurality suffix ‘*...-pi*’. Without this suffix, the reference is 1+2; with the suffix ‘*...-pi*’, the reference is 1+2+3 or 1+3. This is a very interesting referential division, but it is not marked in the pronominal paradigm. The pronominal paradigm has only one form ‘*yk-...*’ for all referential possibilities of ‘we’. The pronominal prefixes from Assiniboine are thus of the ‘unified-we’ type.

Following Greenberg’s analysis, McGregor (1996:40-41) argues that the pronominal prefixes from the Australian language Nyulnyul are also of the ‘Assiniboine-type’. He argued rightly that Nyulnyul resembles Assiniboine strongly, but the consequence is that Nyulnyul does not have a special pronominal paradigm, for the same reason as Assiniboine did not have one. There is only one pronominal prefix for ‘we’ in Nyulnyul, the prefix ‘*ya-...*’. McGregor explains that this prefix is sometimes attested without number marking, in which case the reference is to 1+2.

‘[*ya*] occasionally occurs without the number marking prefix when it refers to the speaker-hearer dyad: that is, when reference is made to the 1&2 minimal category.’ (McGregor, 1996:40)

However, the prefix itself is identical in all referential variations. It is the number suffix ‘*...-rr-*’ that marks the special referential properties of the two ways to say ‘we’ in Nyulnyul. In this case, the problems are even stronger, as the opposition between presence and absence of the number suffix is not completely regular, judging from the example sentences that are given by McGregor (1996). The example presented in (3.6) seems to have 1+2 reference, but there is still a plural suffix in the verb.

(3.6) NYULNYUL

ngay a juy ya-li-rr-jid Derby-ung
 I and you I PLUR-IRR-PLUR-go PLACE-all
 ‘You and I might go to Derby’ (McGregor, 1996:42)

Both the active prefixes from Assiniboine and the pronominal prefixes from Nyulnyul are of the ‘unified-we’ type, following my definitions. The kind of paradigm that Greenberg originally intended to describe as the ‘Assiniboine-type’ exists, but the Assiniboine and Nyulnyul cases are not part of that type. The paradigms attested of ‘real’ Assiniboine-type paradigms will be discussed in section 3.6.6.

3.6.2 Type (Pb): ‘no-we’

Paradigms of type (Pb) are the diametrical opposite of the previous type. Instead of one overarching form for all ‘we’ categories, (Pb) paradigms have no specialised pronominal element for ‘we’ at all. Most examples of pronominal paradigms without any form for ‘we’ do even not have any specialised group marking at all. This is, for example, found in the English inflectional marking, where all the plural forms are identical to the first and second person singular; they are all zero. The English person inflection is a pronominal paradigm of type (Pb). I will use the name ‘no-we’ alternatively instead of (Pb) for this type of paradigmatic structure.

	(Pa)	(Pb)	(Pc)	(Pd)	(Pe)
<i>I+2</i>			A	A	A
<i>I+2+3</i>	A	–	A	A	B
<i>I+3</i>			–	B	C

Table 3.4: the ‘no-we’ patterns (Pb)

Pronominal paradigms of type (Pb), with no specialised marking for ‘we’, are less well known than the other common types of paradigms. It has even been claimed that ‘we’ is a universally grammaticalised category for human languages. This claim implies that the ‘no-we’ type should not exist. The most explicit formulation of this claim is found in Ingram (1978):¹⁸

¹⁸ The claim for the universal existence of the first person plural has been around for a long time. An early hint, and implicit explanation, to this claim is made by Schmidt (1926):

‘Denn gerade bei der Erfassung der Person, besonders der ersten Person, tritt der Begriff der individuellen Einheit am stärksten hervor. Gerade die scharfe Erfassung der Einheit ist aber auch die notwendige psychologische Vorbedingung für die Herausarbeitung der Mehrheitsformen.’ (Schmidt, 1926:316)

Forchheimer (1953), who builds on the work of Father Schmidt, transforms this into a true claim of a universal first person plural:

‘My research has borne out Schmidt’s statement for the plural without exception ... I have found several languages where the word for ‘I’ can also serve to express ‘we’, they all possess, besides that, a word for ‘we’. The only exception is Chinese Pidgin English. ... I found no record of a language distinguishing ‘thou’ and ‘you’, but not ‘I’ and ‘we’, whereas the opposite is frequent.’ (Forchheimer, 1953:12)

A comparable claim, though less strong, returns as Greenberg’s Universal 42: All languages have pronominal categories involving at least three persons and two numbers (Greenberg, 1963:96). Moravcsik (1978) remarks on Greenberg’s universal that it should be interpreted as meaning that the first person universally shows a plural:

[continued on next page]

‘Universal 1: there are at least four persons in every language: I, thou, he, we.’ (Ingram, 1978:227)

Note that in this universal, Ingram talks about ‘languages’. He claims every language has a forms for ‘we’. In contrast, I look at the structure of individual paradigms in a language. In fact, there are two different versions of the universal claim. A strong and a weak version, that have to be distinguished. The strong version claims that every pronominal paradigm has at least four persons:

(3.7) Universal ‘we’, strong version:

There are at least four persons in every PARADIGM: I, thou, he, we.

As the English inflection shows, it is very well possible for an individual paradigm to have no form for ‘we’. Such paradigms are even relatively common among the world’s languages. The complete distribution of this paradigmatic type is discussed in section 4.4. The strong version of the universal ‘we’ is clearly not true. Accordingly, the claim by Ingram must be interpreted as a weaker version of this universal. A possible interpretation is that a particular paradigm may have no form for ‘we, but then there will be another paradigm somewhere in the language that fills this gap. The English inflection may have a ‘no-we’ structure, but English has also independent pronouns with an overt form for ‘we’. This weak version of Ingram’s universal shown in (3.8) holds when all languages that have a ‘no-we’ pronominal paradigm have another pronominal paradigm that includes at least an element for ‘we’.

(3.8) Universal ‘we’, weak version:

There are at least four persons in every LANGUAGE: I, thou, he, we.

However, even in this weaker version, this universal is not absolute. There are a few counterexamples, although they are very rare. The clearest case of a language that contradicts the weak universal is the Brazilian language Mura Pirahã. Mura Pirahã has singular pronouns, and clitics that are derived from the pronouns. The pronouns and clitics can optionally be used co-referentially, as shown in (3.9).

(3.9) MURA PIRAHÃ

<i>ti</i>	<i>gíxai</i>	<i>gí</i>	<i>xibáobá</i>
1	2,PRON	2,CLIT	hit
‘I hit you’			

(Everett, 1987:249)

Besides the pronouns and the related clitics, no other pronominal paradigms exist. The special property of the pronouns (and the clitics) is that they do not have any group-marking elements. They have only elements with singular reference.

‘The three basic Pirahã pronouns are worthy of attention because they comprise one of the simplest pronoun systems known. They are often optional in discourse, meaning that their functional load is not as great as that for pronouns in many languages (especially given the fact that Pirahã has no form of agreement marked on the verb aside from clitics and cliticized pronouns).’ (Everett & Thomason, n.d.)

[continued from previous page]

‘although ... in all languages there will be some context where number distinctions in all three persons will be significant, this does not mean that the personal pronominal paradigm consisting of free forms will have number distinctions manifested in all three persons. The generalization concerning the number distinctions in free personal pronominal forms seems to be this: it is universally present in the first person but not in the second and third.’ (Moravcsik, 1978:352)

There are no markers of number anywhere in the language. There is simply no way to refer overtly to a group of more persons in a grammaticalised form. Other constructions have to be used.

‘Pirahã not only has no plural pronouns, it has no number anywhere in its grammatical system – no number agreement, no singular-plural distinctions in nominals, and no numerals at all. Moreover, there is no direct evidence of plural forms in any surviving Mura or Pirahã data of any period.’ (Everett & Thomason, n.d.)

The only way to mark groups of participants is by conjunction of pronouns (Everett, 1986:281). This pronominal conjunction is zero marked, as shown in (3.10).

(3.10) MURA PIRAHÃ

ti gíxai pío ahápií
1 2,PRON also go

‘You and I will go (ie we will go)’ (Everett, 1986:281)

Other languages that contradict the weak version of the universal ‘we’ are Classical Chinese and the Yuman language Maricopa. In Classical Chinese, there probably was no number marking at all. In any case, nouns and pronouns knew no distinction between singular and plural forms (Norman, 1988:89). In Maricopa, numerous ways to mark number exist, but number is never marked in a pronominal paradigm (Gordon, 1986:21-23, 58, 90-101). Finally, a last language that contradicts the weak version of the universal is the Siouan language Hidatsa. The pronominal prefixes only distinguish between first, second and third person, and there are no plural forms of these prefixes. There is a plural suffix on the verb, but the reference of the plurality is not fixed. As shown in (3.11), both the subject and the object can be interpreted as plural. As far as can be said from the sources, there are no other pronominal elements in Hidatsa (Robinett, 1955; Matthews, 1965:54-61, 174-176).

(3.11) HIDATSA

wa.-há.-th-aʔa-ru

1-with knife-cut-PLUR-COND

‘When we cut it’, ‘When I cut them’ (Robinett, 1955:171)

3.6.3 Type (Pc): ‘only-inclusive’

The next two pronominal paradigms, type (Pc) and (Pd), both distinguish between an inclusive ‘we’ – the categories 1+2 and 1+2+3 combined – and an exclusive ‘we’ – the category 1+3. The central difference between the inclusive and the exclusive is the status of the addressee. The inclusive ‘we’ includes the addressee; there is a ‘2’ in both categories. The exclusive, oppositely, excludes the addressee; it means something like ‘we, I and people associated with me outside the present speech-act’. In the case of paradigms of type (Pc), the exclusive ‘we’ does not have a specialised morpheme. In most cases, though not all, the exclusive ‘we’ is marked by the same morpheme that is used for the first person singular. I will use the mnemonic name ‘only-inclusive’ for this pattern of non-singular marking. The complete list of examples of this type will be presented in section 4.5.

	(Pa)	(Pb)	(Pc)	(Pd)	(Pe)
1+2			A	A	A
1+2+3	A	–	A	A	B
1+3			–	B	C

Table 3.5: The ‘only-inclusive’ pattern (Pc)

This pattern is exemplified here with sentences from Maká, a language from Paraguay. Maká has pronominal prefixes that show a large range of allophonic variation (Gerzenstein, 1994: 98). In the examples presented here, the prefix *hoy*-... is used both for the first person singular, as in (3.12 a), and for the exclusive first person 1+3, as in (3.12 b). The two meanings are distinguished by a plural suffix ...-*ĩt*. The important point here is that there is no specialised morpheme for the marking of the category 1+3 within the pronominal prefixes. The category 1+3 is marked with the same prefix as is used for the first person singular.

(3.12) MAKÁ

- a. *hoy-otoy*
 I-dance
 ‘Yo bailo’ (I am dancing) (Gerzenstein, 1994: 106)
- b. *hoy-otoy-ĩt* *tse-kheen*
 I-dance-PLUR FEM-DEM
 ‘Yo y ella bailamos’ (We -she and I- are dancing) (Gerzenstein, 1994: 103)

By way of contrast, the inclusive ‘we’ is marked by a specialised prefix. The prefix *xi(t)*-... is used for the reference to the speech-act dyad ‘I and you’, as shown in (3.13 a). The same prefix can also be used together with a plural suffix, as shown in (3.13 b).¹⁹ With this plural suffix, the reference is to a group of type 1+2+3: speaker, hearer and others also. The categories 1+2 and 1+2+3 are combined into one ‘inclusive’ pronominal prefix, only to be distinguished by a plural suffix.

(3.13) MAKÁ

- a. *akha’* *xit-otoy*
 2PRON INCL-dance
 ‘Tú y yo bailamos’ (We -you and I- are dancing) (Gerzenstein, 1994: 103)
- b. *inekhwel* *xi-yi-lan-xu’* *na’* *qametenaX*
 1+2+3PRON INCL-3-kill-PLUR DEM tiger
 ‘A nosotros, el tigre nos mata’ (The tiger killed us) (Gerzenstein, 1994: 176)

3.6.4 Type (Pd): ‘inclusive/exclusive’

Paradigms of type (Pd) also distinguish inclusive ‘we’ from exclusive ‘we’ in the same way as in the previous ‘only-inclusive’ (Pc) pattern. However, in the case of (Pd) both categories are marked by specialised morphemes. This pattern will be re-

¹⁹ The referents of *xi(t)*-... in example (3.13b) are in an object role. For participants in an object role a different plural suffix is used as for participants in a subject role; cf this sentence to the sentence (3.12b) above (Gerzenstein, 1994: 138).

ferred to as ‘inclusive/exclusive’. It is an extremely common pattern, with lots of variation. The complete variability will be laid out in section 4.6.

	(Pa)	(Pb)	(Pc)	(Pd)	(Pe)
1+2	A	–	A	A	A
1+2+3			–	–	B
1+3				B	C

Table 3.6: The ‘inclusive/exclusive’ pattern (Pd)

This pattern is exemplified here with some examples from Apalai, a Carib language spoken in Brazil. A specialised inclusive pronominal prefix *s(y)-...* is found in the subject prefixes of Apalai. The two example sentences in (3.14) show that this prefix can be both used for only the present speaker and addressee –ie 1+2; see (3.14a)– and for reference to a general inclusive –ie 1+2+3; see (3.14b); *syt-...* being an allomorph of *s(y)-...*

(3.14) APALAI

- a. *s-yt̃-tase* *kokoro* *j-epe,* *tykase*
 INCL-poison fish-IMP tomorrow 1-friend,POSS say
 ‘“Let us poison fish tomorrow, my friend”, he said’
 (Koehn & Koehn, 1986: 69)

- b. *otoko* *xixi* *a-htao* *eramã̃ko* *ropa* *syt-a-tose*
 where sun COP-when turn back again INCL-COP-PRES,PLUR
 ‘When will we all turn back?’
 (Koehn & Koehn, 1986: 59)

The inclusive prefix *s(y)-...* is different from the exclusive prefix *ynan(y)-...*, as shown in (3.15a). The exclusive prefix is also a specialised non-singular prefix, which is the essential difference between these prefixes from Apalai and the prefixes from Maká in the previous section. The reference to a speaker in Apalai is marked by a prefix *y-...*, which might historically be related to *ynan(y)-...*, but synchronically it is a clearly different prefix.

(3.15) APALAI

- a. *mame* *more* *ynan-urumekane* *ropa* *maikuito-hpe* *exiryke*
 then that EXCL-leave again ant-infested place
 ‘Then we abandoned that place again because of its
 being infested with ants’
 (Koehn & Koehn, 1986: 78)

- b. *y-pipohno*
 1-hit
 ‘I hit it’
 (Koehn & Koehn, 1986: 108)

3.6.5 Type (Pe): ‘minimal/augmented’

The last common type, (Pe), distinguishes three different forms for ‘we’. The inclusive ‘we’ from the previous section is divided into two different forms, called a ‘minimal inclusive’ and an ‘augmented inclusive’. The minimal inclusive refers only to the speech-act dyad of speaker and addressee; ie 1+2. The augmented inclusive refers to any group of participants that at least includes the present speaker and ad-

dressee, but possibly also others; ie 1+2+3. This opposition is added on top of the opposition between an inclusive and an exclusive ‘we’, as in the previous type (Pd). The resulting threefold distinction will be named a ‘minimal/augmented’ paradigm.

	(Pa)	(Pb)	(Pc)	(Pd)	(Pe)
1+2			A	A	A
1+2+3	A	–	A	A	B
1+3			–	B	C

Table 3.7: The ‘minimal/augmented’ pattern (Pe)

As a general characteristic of the description of such a paradigm in a grammar, the meaning of the three different forms for ‘we’ is described in terms like ‘dual’, ‘inclusive’ and ‘exclusive’. Many different approaches to describe the three forms can be found in the literature. Notwithstanding the plethora of descriptive terminology, I have decided to take all these paradigms together and consider them as one and the same type. There are two main criteria for paradigms to be included within this type. First, all pronominal paradigms of this category have three different forms for ‘we’ (which distinguishes this type from the other types that are discussed in this chapter) and, second, these pronominal paradigms do not have any dual forms for the second or third person (which distinguishes this type from paradigms with duals as discussed in section 8.4). It then turns out that this seemingly exotic paradigmatic structure is found relatively frequent among the world’s languages. A full description of the known cases of such paradigms will be presented in section 4.7. In the rest of this section, I will discuss some historical developments that led to the present analysis of this kind of paradigm.

Franz Boas (1911:39) – as quoted in section 3.4 above – first acknowledged the possibility of these three different forms for ‘we’, but he did not know of any language that actually has all three as different morphemes. The meaning of 1+2 and 1+2+3, Boas says, may differ in reference but are not expressed by separate morphemes in any language. If this were true, this would mean that the minimal/augmented type (Pe) does not exist as a distinct type of pronominal paradigm.

‘I do not know of any language expressing in a separate form the combination of the three persons [1+2+3], probably because this idea readily coalesces with the idea of self and person spoken to [1+2].’ (Boas, 1911b:39, number notation added, MC)

However, the distinction between 1+2 and 1+2+3 is found among the world’s languages. The oldest description where this distinction is explicitly noted, is the description by Foster & Foster (1948) of the Mixe-Zoque language Sierra Popoluca in Mexico:²⁰

‘The exclusive plural excludes the person or persons addressed [1+3]. The limited inclusive plural includes the speaker and the person or persons addressed, and excludes

²⁰ Although this quote from the grammar of Sierra Popoluca describes exactly a division of type (Pe), I will not classify this language as belonging to this type. The reason is that the pronominal elements of Sierra Popoluca consist of pronominal prefixes and number suffixes. Following the definitions that I use, the pronominal prefixes are classified as ‘only-inclusive’ (Pc), with separate number marking. See section 4.5.2 for a discussion.

any other who may be present or referred to [1+2]. The generalized inclusive plural includes the speaker, person or persons addressed, and any other person or person present, or absent and referred to [1+2+3].’ (Foster & Foster, 1948:19, number-notation added, M.C.)

The idea of this ‘generalised’ inclusive only caught on in the scientific literature in the wake of a short article by Thomas (1955) on the pronouns in the Philippine language Ilocano. He argued that the traditional description of the Ilocano pronouns is not satisfactory. The traditional description of Ilocano, shown in Table 3.8, distinguishes between an inclusive and an exclusive plural, but also adds a special form for the dual (Bloomfield, 1942:194). The default description of such a language with three different forms for ‘we’ is a division by the terms ‘inclusive’, ‘exclusive’ and ‘dual’. Thomas formulates two points of criticism on this traditional analysis. These two problems will be discussed here shortly. First, the referential properties of the ‘dual’ in these paradigms are more restricted than the name implies and, second, there is no other dual around somewhere else in the language.

‘The 1st dual label is not completely accurate, as the use is restricted to cooperative action by one speaker and one hearer; no one else may be included. ... There is no substantiation from the structure of the rest of the language for the existence of a dual number in Ilocano’ (Thomas, 1955:205).

	<i>Singular</i>	<i>Dual</i>	<i>Plural</i>
<i>1 Inclusive</i>	co	ta	tayo
<i>1 Exclusive</i>			mi
2	mo	yo	
3	na	da	

Table 3.8: Traditional analysis of Ilocano pronouns ²¹

To clarify the first of these problems, I discuss the account of a comparable pronominal paradigm from a recently published grammar of Limbum, a Grassfields language from Cameroon. In the grammar the following description is found of the three forms for ‘we’:

‘Limbum differentiates between exclusive, inclusive and dual ‘we’. The form *wir* is used when the speaker wants to exclude the hearer(s) and *sii* when the speaker wants to include the hearer(s). When there are only two people involved, i.e. the speaker and the addressee, the dual form *sò* is used.’ (Fransen, 1995:179)

In the quote, an ‘inclusive’, an ‘exclusive’ and a ‘dual’ form are mentioned. This is widely known terminology, and that is probably the reason why this terminology is still found abundantly in the descriptive literature. Intuitively, a dual form refers to a group of two persons, but in the quoted description it can be read that the ‘dual’ necessarily includes the addressee. This ‘dual’ is not used for a pair in which the addressee is excluded. The aspect of duality is not a complete description of the meaning of the Limbum ‘dual’. The reference of the form is more specific than that. It is better called a ‘dual inclusive’ (see also Table 3.9).

²¹ In this discussion of Ilocano, the third person pronouns are troublesome. Thomas (1955) argues that they are not real pronouns, but more like number markers. I include them here as if they were normal third person pronouns to preclude confusion; cf section 4.7.3.

	<i>Intuitive division</i>	<i>Actual division</i>
<i>1+3, more than two</i>	‘exclusive’	‘exclusive’
<i>1+3, exactly two</i>	‘dual’	
<i>1+2</i>		
<i>1+2+3</i>	‘inclusive’	‘inclusive’

Table 3.9: Different interpretations of ‘dual-inclusive-exclusive’ in Limbum

The second, and most important, argument against the traditional terminological trinity ‘dual-inclusive-exclusive’ is that the paradigms that distinguish these three forms of ‘we’ normally do not have any other dual forms in other parts of the language.²² These paradigms all look like the example paradigm from Ilocano above. There are no other duals, except for the ‘dual inclusive’.²³ For Ilocano and Limbum, and many others, the pronominal ‘dual inclusive’ is the only dual form in the language. To invoke a completely new category – the dual – for the analysis of one morpheme is a waste of apparatus. In the wake of Thomas’ ground breaking analysis for Ilocano, Conklin adds, slightly ironical, that the traditional analysis is ‘hardly elegant, economical, or convincing’ (1962:134). A better approach to paradigms with three different forms for ‘we’ is to focus on the KIND of participants, instead of the NUMBER of participants. Going back to the Ilocano pronouns, the opposition between the dual *ta* and the inclusive *tayo* is better labelled by different terms. Using the terminology as developed in section 3.4, the opposition can be expressed as a difference between including only the addressee in the case of *ta* (ie 1+2) and including the addressee and others in the case of *tayo* (ie 1+2+3).

In the original article, Thomas (1955) generalised the opposition between 1+2 and 1+2+3 to the whole paradigm, as shown in Table 3.10. He aligned 1+2 with the singular persons and analysed 1+2+3 as the ‘plural’ of 1+2. As the terms ‘singular’ and ‘plural’ are obviously inadequate, he used the terms ‘simple’ vs. ‘more’:

‘... a third analysis ... would present the oppositions of Speaker : Hearer : Speaker-Hearer persons and simple : «more» or «plus» numbers. This «more» component indicates that there is more than just the primary person.’ (Thomas, 1955:208)

Conklin (1962:135) introduced the terms ‘minimal membership’ versus ‘nonminimal membership’ instead of ‘simple’ vs. ‘more’. Today, this phenomenon is normally re-

²² This claim is substantiated by the data as presented in table 11 from Plank (1996:131). In the large database on various forms of dual marking that has been compiled by Plank and his co-workers, there is only one language that combines the marking of a ‘first person inclusive dual only’ with nominal dual marking. The special status of the first person pronominal marking for the occurrence of dual forms was already noted by von Humboldt (1827:156) and later reinforced by Plank (1989:305,311-312; 1996:130-131). Humboldt also presents an explanation for the preference of a dual in the first person: the combination of speaker and addressee form a ‘natural pair’.

²³ There are a few pronominal paradigms among the world’s languages that distinguish three ‘we’-categories and also have a dual for the second and third person. These paradigms are, consequently, interpreted as a different group. They will be discussed in the chapter on dual marking, in section 8.4.

ferred to as an opposition ‘minimal’ vs. ‘augmented’, following McKay (1978).²⁴ Taken at face value, this analysis presents a nice coherent picture of the paradigmatic structure. Surely, when compared to the traditional analysis, as shown in Table 3.8 above, the minimal/augmented analysis present an impressive improvement. Still, there are two problems with this analysis. Both problems concern the generalisation of the opposition between 1+2 and 1+2+3 to the whole paradigm, not the opposition between 1+2 and 1+2+3 as such.

	<i>Minimal</i>	<i>Augmented</i>		
1+2	ta	tayo	1+2	+3
1	co	mi	1	+3
2	mo	yo	2	+3
3	na	da	3	+3

Table 3.10: Minimal/augmented analysis of Ilocano pronouns

The first problem is that the MORPHOLOGICAL structure of these paradigms does not confirm the minimal/augmented analysis. If the minimal/augmented opposition were valid for the whole paradigm, then one would expect to find at least some pronominal paradigms with morphologically transparent marking of the four augmented categories. In such a paradigm, the four augmented categories would be derived morphologically from the four minimal categories by way of a regular ‘augmented’ morpheme. However, I have never come across such a morphologically transparent minimal/augmented paradigm among the world’s languages. The second problem concerns the HISTORICAL CONNECTION of the minimal/augmented paradigm to other paradigms. The generalised analysis as shown in Table 3.10 seems to argue for a connection to the ‘only-inclusive’ type (Pc). The only-inclusive paradigm looks like the minimal set of a minimal/augmented paradigm, and a historical development from the only-inclusive of the minimal/augmented paradigm seems a reasonable guess. However, it turns out to be impossible to find any good evidence for this transition.²⁵ There is much better evidence for the transition from the ‘inclusive/exclusive’ type (Pd) to the minimal/augmented paradigm (cf Greenberg, 1988:6). The full evidence for this transition will be presented in section 6.4.2. Anticipating that discussion, the connection to inclusive/exclusive (Pd) paradigms can be made plausible by looking at the Ilocano paradigm. The pronoun *tayo* in Ilocano is the odd one out as it is the only one which is disyllabic. It looks like *tayo* is a later addition, made up from the 1+2 pronoun *ta* and the 2+3 pronouns *yo*. Interestingly, in the closely related language Pampangan, the equivalent pronoun *tami* is also the only disyllabic pronoun, but in

²⁴ As a result of this analysis, there is a small current in the descriptive literature to call the minimal inclusive a ‘fourth’ person, as this inclusive goes along with the other ‘singular’ marking (eg Hardman, 1966:56; Osborne, 1974:38; Gerzenstein, 1994:83). This usage of ‘fourth’ person is semantically awkward and conceptually confusing. It should be avoided (cf Hymes, 1972:105).

²⁵ There are many examples of only-inclusive (Pc) paradigms that can be pluralised. In these cases, the plural forms have the same referential values as an augmented set. However, these plurals are never marked directly in the pronominal system. In all cases I have seen, the plural marker for a (Pc) paradigm is marked independently from the pronominal elements, eg the plural marker is a suffix, while the pronominal elements are prefixes.

this case it is made up from the 1+2 pronoun *ta* and the 2 singular pronoun *mi* (Gonzales, 1981:172-173). The two different structures of this pronoun argues for a late addition. Also, there are many Philippine languages that have an inclusive/exclusive (Pd) pattern (cf Reid, 1971), but I know of none that has an only-inclusive pattern (Pc). For these reasons, paradigms like the Ilocano pronouns will be presented in yet a different analysis in the present work. The paradigms will be displayed with two columns, distinguishing the singular and the group morphemes, as shown in Table 3.11.²⁶

		<i>'singular'</i>		<i>'group'</i>	
				ta	1+2
				tayo	1+2+3
1	co			mi	1+3
2	mo			yo	2+3
3	na			da	3+3

Table 3.11: Singular-Group analysis of Ilocano pronouns

3.6.6 Rare types

Finally, five other patterns of the first person complex are attested among the world's languages, but these patterns are extremely rare. The examples of these patterns are incidental cases, probably only to be explained separately for each case with its own peculiar history. I will only synchronically describe all these rare cases in this section, and disregard them in later chapters. The structures of the first person complex are repeated in Table 3.12.

		<i>(Pf)</i>	<i>(Pg)</i>	<i>(Ph)</i>	<i>(Pi)</i>	<i>(Pj)</i>
1+2	A	A		A	-	
1+2+3	B	B		B	A	-
1+3	B	A		-	A	A

Table 3.12: Rare patterns of the first person complex

These five structures seem to fall in two classes. The first two patterns (Pf) and (Pg) do not have any overlap between singular and non-singular categories, and these two are both attested in more than one case. In the other three patterns, an overlap between singular and non-singular is found and these three patterns are each represented by only one case. However, the number of examples is so small that it is difficult to judge whether this distinction is typologically salient. Future research has to bear out

²⁶ Forchheimer (1953) seems to hint at a distinction between the minimal/augmented analysis (cf Table 3.10) and the singular-group analyses (cf Table 3.11) on the basis of the morphological structure of the forms for 'we'. His type IV-A4 can be compared to the 'singular-group' analysis, and his IV-A2 to the 'minimal/augmented' analysis (ibid.:83-84). However, I do not see any difference that is crucial in this respect between the presented paradigms from Winnibago (ibid.:88) and Sierra Popoluca (ibid.:93).

whether the first two patterns have a different status compared to the others. The unusual structures of the pronominal paradigms that will figure in this section will be described by using the scheme as developed in Table 3.11. The first column shows the singular forms, the second column will show the plural forms.

The first of these rare types, paradigms of type (Pf), has two different forms for ‘we’. Taken at face value, this might resemble an ‘inclusive/exclusive’ type (Pd) paradigm, but actually the referential values of the two morphemes are different. In the case of (Pf), there is a specialised morpheme for the speech-act dyad and another specialised morpheme for all other references of ‘we’. This second morpheme includes all exclusive reference, but also some inclusive reference.²⁷ I know of three cases of this kind of paradigms, the first two of which are inflectional. In both these inflectional cases, the referential values of the inflectional morphemes can be exactly established because the independent pronouns in both languages distinguish between 1+2+3 from 1+3. The two different pronouns are cross-referenced by the same affix on the verb. Shown in (3.16) is the first example of this unusual paradigmatic structure. This is the paradigm of the pronominal prefixes from the Australian language Bardi (Metcalf, 1975: 123).²⁸

(3.16) BARDI

		‘group’		
		a-...		1+2
	‘singular’	aŋ-...		1+2+3
1	ŋa-...	guŋ-...		1+3
2	mi-...	iŋ-...		2+3
3	i-...			3+3

Shown in (3.17) is the second inflectional example of (Pf): the paradigm of the imperfect suffixes from the Papuan language Kunimaipa. In the case of Kunimaipa, other inflectional paradigms in the language are attested with different structures, one of which will follow shortly. The contrast between the various paradigmatic structures ensures that the strange looking structures are genuine (cf Pence, 1968: 110; Geary, 1977: 26).²⁹ The final example of type (Pf) is the paradigm of the independent pronouns of the Siouan language Lakota (Van Valin, 1977: 74-75).³⁰

²⁷ Greenberg (1988:9) uses the name ‘Assiniboine-type’ for these paradigms, after the Siouan language Assiniboine. However, the case of Assiniboine is strictly speaking not of type (Pf) by my definitions. Assiniboine person marking consists of pronominal PREFIXES of the ‘unified-we’ type (Pa) with separate number SUFFIXES. The number suffixes are used in a special way, turning the prefixes into a (Pf)-like paradigm. However, by my definitions, only the prefixes are included in the typology. See section 3.6.1 for an extensive discussion of the Assiniboine prefixes.

²⁸ The independent pronouns from Bardi make up a minimal/augmented paradigm. They can be found in Metcalf (1975:49-50,129).

²⁹ The independent pronouns from Kunimaipa have a unusual structure involving also dual forms. See section 8.6 for a discussion of these independent pronouns from Kunimaipa.

³⁰ The independent pronouns from Northern Sierra Miwok and the neighbouring Plains Miwok might also belong to this (Pf) type of paradigm. The inclusive pronoun is translated as meaning ‘I and you’, indicating that it might be only used for the dual inclusive 1+2 (cf section 9.3). However, it is not completely clear whether it is exclusively used for the speaker-addressee dyad, or whether it might

(3.17) KUNIMAIPA

		‘group’		
			...-paine	1+2
	‘singular’		...-ka	1+2+3
1	...-ma		...-pike	1+3
2	...-ke			2+3
3	...-pa(ne)			3+3

(3.18) LAKHOTA

		‘group’		
			ũkiye’	1+2
	‘singular’		ũkiye’pi	1+2+3
1	miye’		niye’pi	1+3
2	niye’		iyē’pi	2+3
3	iyē’			3+3

Paradigms of type (Pg) distinguish a separate form for 1+2+3, but unite the references to 1+2 and 1+3. This paradigmatic structure is found in the subject pronouns from Yaouré, a Mande language from Ivory Coast. There are two different forms of ‘we’ in Yaouré. The special morpheme for 1+2+3 is the odd one out, as its morphophonological behaviour is different from the other pronouns. Possibly, this 1+2+3 pronoun is a recent addition to the pronominal paradigm.³¹

‘kàà, which has an inclusive reference (the speaker and a group of listeners) and kŭ, which has either a dual reference (you and I) or an exclusive reference (the others and I). ... The kàà form of the first person plural frequently reacts differently from the other pronouns.’ (Hopkins, 1986:192)

(3.19) YAOURÉ

		‘group’		
			kŭ	1+2
	‘singular’		kàà	1+2+3
1	ã		kŭ	1+3
2	ĩ		kā	2+3
3	ē		ō	3+3

Another example of a (Pg) type paradigm is found in the Australian language Gooniyandi. The independent pronouns with a (Pi) structure are shown in (3.20). The difference between the two possible meanings of *ngidi* can be disambiguated by number

[continued from previous page]

also be used in other, less prototypical, inclusive reference (Callaghan, 1974:386; 1984:296-297; 1987:397-398).

³¹ In Dan, another Mande language from Ivory Coast, there is a full 8-way pronominal system, with distinct elements for the categories 1+2, 1+3 and 1+2+3 (Doneux, 1968:45-47). This indicates that the Yaouré structure is probably an exceptional structure, not only exceptionally world-wide, but also exceptionally within its own family.

suffixes. However, these number suffixes are only optionally used (McGregor, 1989:438-439; 1990:167-169).

(3.20) GOONIYANDI

		'group'		
		ngidi		1+2
	'singular'	yaadi		1+2+3
1	nganyi	ngidi		1+3
2	nginyi	gidi		2+3
3	niyi	bidi		3+3

A paradigm of type (Ph) separates the two different forms of the inclusive 'we' (1+2 and 1+2+3), already a rather exotic thing to do from a West European perspective. Moreover, the paradigm fails to explicitly refer to the exclusive 'we' 1+3. The only example of a paradigm of type (Ph) that I am aware of comes from the Australian language Tiwi. The subject prefixes from Tiwi come in different variants. The version that is of interest here is the paradigm that is used when there is a masculine third person object, and the tense is non-past. This paradigm is shown in (3.21) below. Other prefixal paradigms from Tiwi do not show this structure. Even within the language itself, this seems to be a marked paradigmatic structure (Osborne, 1974:38).

(3.21) TIWI

		'group'		
		mu-...		1+2
	'singular'	ŋa-...		1+2+3
1	ŋə-...			1+3
2	ŋə-...			2+3
3	a-...	wu-...		3+3

The perfective suffixes from Kunimaipa have a different structure compared to the imperfective suffixes that were shown in (3.17). The perfective suffixes, shown in (3.22), are the only example of a paradigm of type (Pi) that I am aware of. In this paradigm, the 1+2 reference is made by the same morpheme as is used to refer to the speaker. The other referential possibilities of 'we' are taken together with the second person plural (Pence, 1968: 110).

(3.22) KUNIMAIPA

		'group'		
		...-ho		1+2
	'singular'	...-gi		1+2+3
1	...-ho	...-gi		1+3
2	...-ngi	...-gi		2+3
3	...-ha			3+3

A pronominal paradigm of type (Pj) shows the inverted structure of the 'only-inclusive type (Pc). Both (Pj) and (Pc) have only a single grammaticalised form for

‘we’, but the meaning of this ‘we’ is exactly the opposite. In the case of (Pc), the special ‘we’-form has an inclusive reference: combining 1+2 and 1+2+3 but excluding 1+3. In the case of (Pj), the special ‘we’-form has the opposite referential value. The morpheme for ‘we’ refers to a group of category 1+3, and it is not used for reference to 1+2 or 1+2+3. One of the singular forms is used for the reference of 1+2 and 1+2+3. I know of only one example of this type (Pj), found in the inflection of the Papuan language Binandere. Shown in (3.23) are the ‘past II stative’ suffixes, but all other suffixal paradigms have the same structure (Capell, 1969b: 16-31).

(3.23) BINANDERE

		<i>‘group’</i>	
		...-ana	1+2
	<i>‘singular’</i>		1+2+3
1	...-ana	...-ara	1+3
2	...-ata	...-awa	2+3
3	...-evira	...-ara	3+3

In general, there seems to be a slight preference for the inflectional marking among these rare patterns. However, the number of examples is so small and the structures are so heterogeneous that it is difficult to draw any conclusive generalisations about this group. The main point that can be made is that these structures are unusual for a human language.

3.7 Generalisations

Five patterns out of the 15 theoretical possibilities for the first person complex are common among the world’s languages. These five patterns are repeated here in Table 3.13. The only paradigms attested that do not belong to one of these five types were discussed in the previous section 3.6.6. The relative frequency of the five common patterns will be discussed extensively in the next chapter. In this section, I will formulate some preliminary generalisations on the structure of the marking of the first person complex. These generalisations will separate the five common patterns from the rare or unattested patterns. These generalisation will be further refined in chapter 5, after the discussion of all the details of the variation attested in chapter 4.

	<i>(Pa)</i>	<i>(Pb)</i>	<i>(Pc)</i>	<i>(Pd)</i>	<i>(Pe)</i>
1+2	A	–	A	A	A
1+2+3			–	B	C
1+3					

Table 3.13: Common patterns of first person complex

In the majority of the world’s languages, there is no separation between 1+2 and 1+2+3. Together these two categories form the ‘inclusive’ category, as opposed to the ‘exclusive’ 1+3. In Table 3.14 all theoretical possible marking patterns with only these two referential clusters are shown. These five possibilities are indeed all attested, although the last pattern (Pj) is one of the rare patterns as described in section

3.6.6. This pattern is possible, but not probable for a pronominal paradigm. The other four cases are common.

	(Pa)	(Pb)	(Pc)	(Pd)	(Pj)
'Inclusive' (1+2, 1+2+3)	A	–	A	A	–
'Exclusive' (1+3)			–	B	A

Table 3.14: Inclusive-exclusive marking patterns

When the single example of (Pj) is disregarded, an interesting generalisation arises. A specialised inclusive can exist without a specialised exclusive, but not vice versa. An exclusive can not exist without an inclusive.³² When the types (Pa) and (Pb) are taken together – both are cases with no opposition between an inclusive and an exclusive – then this generalisation can also be formulated as an implication:

(3.24) **Addressee Inclusion Implication I**³³

Exclusive → Inclusive

The argument behind this implication is illustrated by the cross-tabulation in Table 3.15. In words: if there is specialised marking for the exclusive in a pronominal paradigm, then there is specialised marking for the inclusive. Or formulated in reverse order: there can only be a specialised exclusive when there is already a specialised inclusive.

		Specialised Inclusive	
		Yes	No
Specialised	Yes	(Pd)	–
Exclusive	No	(Pc)	(Pa), (Pb)

Table 3.15: Implication 'exclusive → inclusive'

Next, the inclusive can be divided into two categories: a minimal inclusive (1+2) and an augmented inclusive (1+2+3). The split between the categories 1+2 and 1+2+3 is almost exclusively found in the form of a paradigm of the minimal/augmented type (Pe). There are only a few incidental cases of differently structured paradigms with a split between 1+2 and 1+2+3 (see section 3.6.6). These rare variants will be disregarded here. The first generalisation on such split inclusives is that both parts (1+2 and 1+2+3) are overtly marked. More interestingly, in case of a split inclusive, there is always a specialised exclusive. This generalisation can also be formulated as an implication:

(3.25) **Addressee Inclusion Implication II**

Split Inclusive → Exclusive

³² With 'specialised' I mean that there is an overt morpheme that specifically codes for this function and not for any other functions in the pronominal paradigm. For example, a 'specialised inclusive' is a non-zero morpheme that only marks the inclusive categories 1+2 and 1+2+3, nothing more, nothing less.

³³ Cf Universal 43 from Sokolovskaya (1980:95).

The rationale behind this implication is illustrated by the occurrences as shown in Table 3.16. In words: if there is a split inclusive in a pronominal paradigm, then there is specialised marking for the exclusive 1+3. Or formulated in reverse order: there can only be a split inclusive when there is already a specialised exclusive.

		Split Inclusive	
		Yes	No
Specialised	Yes	(Pe)	(Pd)
	Exclusive	No	(Pa), (Pb), (Pc)

Table 3.16: Implication ‘split inclusive → exclusive’

These two implications can be combined into a hierarchy, but some care has to be taken in the formulation of this hierarchy. It might be tempting to concatenate the two implications to something like ‘split inclusive → exclusive → inclusive’, but that is logically not a valid generalisation. The implications can be combined into a hierarchy when they are read as conditions. This is clear when both ‘reversed’ formulations, as I have called them above, are taken together. These ‘reversed’ implications are repeated here as (3.26):

(3.26) **Reversed Addressee Inclusions Implications**

- there can only be an exclusive when there is also an inclusive;
- there can only be a split inclusive when there is also an exclusive.

Combined, this means that there has to be an inclusive for there to be an exclusive, and there has to be an exclusive for there to be a split inclusive. This hierarchy of conditions is set out as a hierarchy of questions in Table 3.17.

<i>Is there any specialised form for ‘we’?</i>					
<i>No</i>	<i>Yes</i>				
↓	<i>Is the inclusive specialised?</i>				
	<i>No</i>	<i>Yes</i>			
	↓	<i>Are both the inclusive and the exclusive specialised?</i>			
		<i>No</i>	<i>Yes</i>		
		↓	<i>Is the inclusive split?</i>		
<i>No</i>	<i>Yes</i>				
	↓	↓	↓	↓	
no-we (Pb)	unified-we (Pa)	only-inclusive (Pc)	inclusive/exclusive (Pd)	minimal/augmented (Pe)	

Table 3.17: Hierarchy of types of the first person complex

The first choice in the hierarchy is whether there is any specialised marking for the first person complex at all. If there is, then the next question is whether there is specialised marking for the inclusive. If that is available, the next decision is whether

both the inclusive and the exclusive are overtly marked by a specialised morpheme. Finally, if all of this is the case, then the inclusive can be split into a minimal inclusive and an augmented inclusive. The result is an implicational hierarchy of paradigmatic types. The hierarchy consists of a series of conceptual choices that have to be made. These choices are interdependent. After each choice, only one of the possible decisions leads to a new choice. These choices form a hierarchy, and consequently the resulting types form a hierarchy as well. This hierarchy of types can be summarised by as shown in (3.27).

(3.27) First Person Complex Hierarchy

no-we > unified-we > only-inclusive > inclusive/exclusive > minimal/augmented

3.8 Conclusion

This chapter started with some criticism on the concept ‘plurality’ within the pronominal domain. ‘Plurality’ is semantically a rather awkward term for pronominal elements. ‘Plural’ pronouns are not always plural; more often their meaning is associative. Instead of ‘plural’, the perspective of ‘group’ marking was introduced, and followed throughout the chapter. The group perspective is not only semantically better equipped for the analysis of pronominal elements, it also naturally includes the widespread distinction between inclusive and exclusive ‘we’.

There are seven theoretical possible groups. Five of these seven different categories are actually attested in the world’s linguistic variation. The five categories in the group perspective are: 1+3 (exclusive ‘we’), 1+2 (minimal inclusive ‘we’), 1+2+3 (augmented inclusive ‘we’), 2+3 (plural ‘you’) and 3+3 (‘they’). These five categories can be combined theoretically into 203 different marking patterns. As a start, only the marking patterns of the first person complex were taken up in this chapter. The first person complex consists of the categories that include reference to the speaker: 1+3, 1+2 and 1+2+3. These three categories are all subsumed under the English pronoun ‘we’. The three categories of the first person complex can be combined theoretically into 15 different marking patterns. Of these 15 theoretical patterns, ten are attested in the world’s languages. Only five out of these ten are common, the other five are only found in incidental cases. The five common patterns of the first person complex are described in Table 3.18 (see next page).

A few generalisations can be made on the basis of the structure of these five common cases. There are two implicational universals that can be combined into a kind of hierarchy. The following terminology is used for the formulation of the universals. A specialised morpheme for the category 1+3 is called ‘exclusive’ and a specialised morpheme for the combination 1+2 and 1+2+3 is called ‘inclusive’. If the two inclusive categories are both marked by a separate morpheme, this is called ‘split inclusive’. These two universals are repeated here as (3.28) and (3.29).

(3.28) Exclusive → Inclusive

(3.29) Split Inclusive → Exclusive

<i>Name of type</i>	<i>Description of pattern</i>
<i>unified-we</i>	all three categories (1+2, 1+2+3, 1+3) are marked by one specialised morpheme
<i>no-we</i>	none of the three categories (1+2, 1+2+3, 1+3) is marked by a specialised morpheme
<i>only-inclusive</i>	the categories 1+2 and 1+2+3 are marked by a specialised morpheme; the category 1+3 is not marked by a specialised morpheme
<i>inclusive/exclusive</i>	the categories 1+2 and 1+2+3 are marked by a specialised morpheme; the category 1+3 is marked by a separate specialised morpheme
<i>minimal augmented</i>	All three categories (1+2, 1+2+3, 1+3) are marked by a separate specialised morpheme

Table 3.18: Common marking types of the first person complex

These two universals can be combined into a hierarchy of conditions. There has to be an inclusive for there to be an exclusive, and there has to be an exclusive for there to be a split inclusive. The types that are the result of the different fulfilled conditions are shown in (3.30) as a hierarchy.

(3.30) no-we > unified > only-inclusive > inclusive/exclusive > minimal/augmented

It should be kept in mind that this is only a preliminary version of the definitive model of paradigmatic variation that will be developed in the next chapters. In general outline it is accurate, but some refinements and additions will be necessary once more categories and historical considerations are taken into account. However, I will first turn to a complete description of the paradigmatic variability of the five common patterns of the first person complex.

Part Two

Paradigmatic structure

‘By focusing attention upon a small range of relatively esoteric problems, the paradigm forces scientists to investigate some part of nature in a detail and depth that would otherwise be unimaginable.’

Thomas S. Kuhn, *The structure of scientific revolutions*

The grammatical term ‘paradigm’ comes from the Greek *parádeigma*, which simply means ‘example’. In the course of the first centuries BC, this word acquired the specific meaning of a prototypical exemplar for grammatical declension or inflection, as in the famous series *amo, amas, amat* or *rosa, rosae, rosae, rosam, rosa*. In the modern usage, the meaning of the term ‘paradigm’ refers to a set of grammatically conditioned forms that are all derived from a single root or stem. This sense of the term ‘paradigm’ is attributed to Ferdinand de Saussure. In essence, this attribution is right, yet de Saussure used *rapport associatif* for what is today called a ‘paradigm’, and reserved the word *paradigme* for the classical meaning of declension and inflection (de Saussure, 1916: 15, 170-180). Much later, Thomas Kuhn, in *The structure of scientific revolutions*, adopted the term ‘paradigm’ in imitation of the traditional grammatical usage of the word: ‘In its established usage, a paradigm is an accepted model or pattern, and that aspect of its meaning has enabled me, lacking a better word, to appropriate ‘paradigm’ here’ (Kuhn, 1962:23).

In this Part Two, I will investigate the structure of paradigms (in the modern sense of the word) that are used for person marking. In chapter 4, an extensive typology of the attested paradigmatic structures will be presented. The first goal of this chapter is to present the wide variety of possibilities that is attested among the world’s languages and to classify this variation into a typology. Second, and most importantly, I will categorise the various types on a continuous scale between ‘common’ and ‘rare’, so as to add some relief to the monotone listing of possibilities. In chapter 5, this typology will be analysed quantitatively, resulting in two hierarchies: the Explicitness Hierarchy and the Horizontal Homophony Hierarchy. Finally, in chapter 6, these hierarchies will be put to the diachronic test. If the typological hierarchies make any sense, then diachronic changes are expected to follow roughly along the lines of the hierarchies. This hypothesis will be tested in this chapter.

Chapter 4

The diversity of the core

A survey of possible patterns for singular and group marking

4.1 Introduction

In the classical analysis of the *pronomina*, various *accidentia* ('attributes') of the pronoun were distinguished. For example, in the *Téchnē Grammatikē* by Dionysius Thrax 'persons, genders, numbers, cases, shapes, species' are mentioned as possible attributes of pronouns (Uhlig, 1883:64; Kemp, 1987:182). Somewhat differently, in the *Ars Minor* by Donatus the attributes 'quality, gender, number, form, person, case' are listed (Chase, 1926:33). Many slightly different version of the list of attributes can be found in the classical literature, but all agree on one basic point: there is an *unordered* list of characteristics that are equally relevant to the analysis of pronouns. This classical analysis is rather different from what is found in more recent grammatical work. Today, the *personal* pronouns have been put on top as the most important kind of pronouns, and the attributes 'person' and 'number' have become the central pivot for their description. However, there has never been a clear decision to change the perspective. Over the ages, the special status of the combination person/number has diffused slowly and largely unnoticed into the grammatical practice. It dates back at least to *De Emendata Structura* (1524) by Thomas Linacre. He gives these two attributes a special place among the others because they occur with all declinable word classes (Luhrman, 1984:185,205).¹

Over the following centuries, the combination person/number virtually becomes the basic framework to describe pronouns. Today, the special status of the person/number combination can be found in almost all grammars of particular languages, but also in more general works on personal pronouns, like Forchheimer (1953) and Kuryłowicz (1964:148-157) to name just a few. The general acceptance of the primacy of the person/number-combination indicates that there is a sense in which these two belong together, although a justification is never made explicit. The correctness has always been tacitly assumed. Not everybody, however, agrees on the special status of the combination person/number. Especially, many of the more theoretically oriented works on grammatical structure still treat 'person' as an individual attribute, separate

¹ The same special treatment of the combination person/number is also found in *De Causis* (1540) by Julius Scaliger (Luhrman, 1984:243). The first prolegomena to the prominence of the person/number combination can also be found in early descriptions of English (Vorlat, 1975:199-218) and Dutch (Dibbets, 1995:204-206).

from number (Jespersen, 1924:212-215; Lyons, 1968: 276-281; 1977: 636-646; Croft, 1990: 149-150).

From this short historical survey it can be inferred that there is a sense in which person marking in the singular and in the plural is at the core of pronominal marking. However, at the same time, the dimensions ‘person’ and ‘number’ seem to be two different dimensions. In this chapter, I will present a position in between these two lines of thought. For a proper cross-linguistic comparison, it seems most promising to divide the dimensions ‘person’ and ‘number’ slightly differently from the traditional usage. The dimension ‘number’ is split into two parts: ‘general’ non-singular reference (ie plural) and ‘restricted’ non-singular reference (ie dual, paucal, etc.). ‘General’ non-singular reference is defined as the marking that makes no difference as to the cardinality of the included referents, as long as the number is greater than one. In contrast, for ‘restricted’ non-singular reference, the cardinality of the set of referents is of crucial importance. In this chapter, ‘general’ non-singular is taken together with the singular. This combination forms the set of pronominal elements that is unmarked for number (see Figure 4.1). The discussion of ‘restricted’ non-singular reference will be postponed to Part Three.

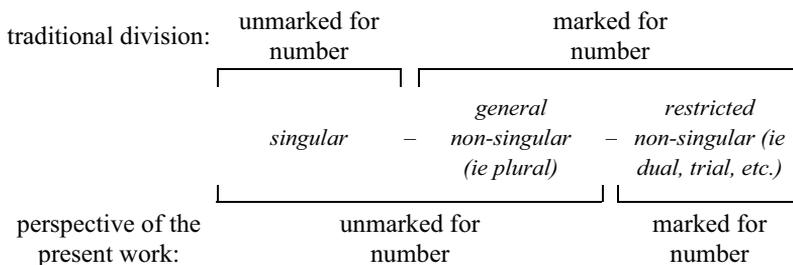


Figure 4.1: Different divisions of the dimension ‘number’ in pronouns

The unity of singular and general non-singular becomes an obvious combination when the notion of ‘plural’ is replaced with the notion of ‘groups’, as argued in the previous chapter. Different kinds of groups of persons were distinguished on the basis of various combinations of the three singular participants. Five kinds of groups turned out to be salient linguistic categories. The three singular persons and the five groups of persons make up a core of eight categories for person reference.² Not all pronominal paradigms grammaticalise all eight referential categories. Some paradigms mark them all, others distinguish only a few different morphemes. In this chapter, I will present the attested diversity of paradigmatic structures, ranging from paradigms with only two different morphemes up to paradigms that maximally distinguishing all eight categories by different morphemes. Much is possible, even when only these eight

² In fact, this means that in the present work the two attributes ‘person’ and ‘number’ are combined by tacitly including ‘number’ under the heading of ‘person’. Interestingly, this is highly similar to what is found in two ancient Dutch grammars: ‘Twe-spraack’ (1584) and ‘Ont-werp der Neder-duitsche letterkonst’ (1649). In both grammars, the attribute ‘number’ is not mentioned in the discussion of pronouns; the plural forms are added to the discussion of the person-categories without using an extra dimension of ‘number’ (Dibbets, 1995:204-205).

‘core’ categories are taken into account. However, not everything is equally likely to be found. The main goal of this chapter is to decide which paradigmatic patterns are common, and which are rare.

This chapter has been set up as follows. In section 4.2, the meta-language for the description of the various paradigms is explained. This meta-language will be used throughout the chapter, casting all examples in the same mould. The result of this unification is that the various structures can easily be compared. Also in section 4.2, the criterion for distinguishing between ‘common’ and ‘rare’ is defined. The following five sections of this chapter will consist of a discussion of the variation within the five major types of the first person complex, as established in the previous chapter. The presentation in these sections 4.3 to 4.7 will be rather lengthy, listing example after example of many different paradigmatic structures. This perhaps somewhat boring part had to be compiled to be able to draw the more interesting conclusions afterwards. The structural analysis of the variation attested can be found in the following two chapters. This chapter concludes with a summary of the attested variation in section 4.8.

4.2 Method of classification

The presentation of the various paradigmatic structures in this chapter is streamlined by using the same meta-language for all paradigms. In this way, the different structures can be easily compared. This meta-language consists of eight different categories: the three singular categories (1, 2 and 3) and the five different groups that were distinguished in the last chapter (1+2, 1+2+3, 1+3, 2+3 and 3+3). This 8-way division is an improvement compared to the traditional 6-way division. First, it avoids the semantically troublesome notion of ‘plural’ in the pronominal domain. Second, it takes care of the inclusive-exclusive distinction in an integrated way, as argued extensively in the preceding chapter. This 8-way division will be used instead of the traditional 6-way division as a reference structure for the discussion of pronominal paradigms. The paradigms will be shown in a fixed graphic format with the singular categories in the leftmost column and the group-categories in the rightmost column. This format is presented in Figure 4.2. The various labels that are used for the different categories are added to the figure.

		<i>‘non-singular’</i>			
<i>‘singular’</i>		1+2	<i>minimal inclusive</i>	} <i>inclusive</i>	} <i>first person complex</i>
		1+2+3	<i>augmented inclusive</i>		
<i>speaker</i>	1	1+3	<i>exclusive</i>		
<i>addressee</i>	2	2+3	<i>second person plural</i>		
<i>other</i>	3	3+3	<i>third person plural</i>		

Figure 4.2: Outline for presentation of paradigmatic structures

Different categories that are marked by one morpheme are shown as contiguous blocks without lines in between the categories. The term ‘homophony’ is used in a

theory-neutral sense to talk about different kinds of combinations of the various categories. I distinguish between three different kinds of homophony. A SINGULAR HOMOPHONY is a combination of different singular categories into one morpheme. The possible variants of singular homophony were already discussed in chapter 2. A VERTICAL HOMOPHONY is a combination of different non-singular categories into one morpheme.³ Finally, a HORIZONTAL HOMOPHONY is a combination of a singular with a non-singular category into one morpheme. These three kinds of homophony are exemplified in Figure 4.3. In this figure, only homophony is shown between categories that happen to be contiguous blocks in the graphic format that is used. This is not necessarily so. Sometimes, non-contiguous categories are homophonous. In such cases, the blocks are connected with small corridors, as can, for example, be seen in examples (4.21) and (4.23) on page 115 below.

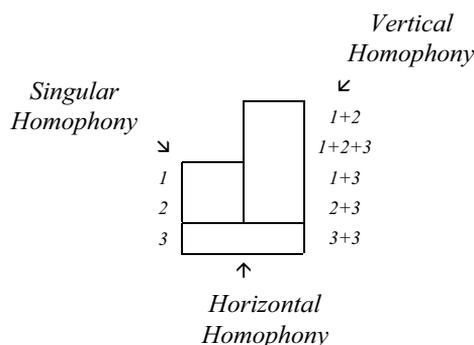


Figure 4.3: Different kinds of homophony

In the previous chapter, I argued that the first person complex is the central classificatory device for the typology of paradigmatic structure. The first person complex is the set of group morphemes that include at least the first person, viz the categories 1+2, 1+2+3 and 1+3 (see also Figure 4.2 above). Five of the 15 possible patterns of the first person complex account for the far majority of the attested paradigmatic variation. These five major patterns of the first person complex are repeated here in Table 4.1.

	<i>unified-we</i>	<i>no-we</i>	<i>only-inclusive</i>	<i>inclusive/exclusive</i>	<i>minimal/augmented</i>
<i>1+2</i>	A	–	A	A	A
<i>1+2+3</i>			–	B	C
<i>1+3</i>			–	B	C

Table 4.1: Major patterns of first person complex

³ More precisely formulated, a vertical homophony is a combination between categories from the first person complex and the second or third person plural into the reference of one morpheme. Another possible combination that is also classified as a vertical homophony is the combination of the second and third person plural. The various combinations of the three different kinds of ‘we’ into the marking of one morpheme are not considered to be cases of vertical homophony.

Other structures of the first person complex were also attested, but these turned out to be only incidental cases. To give an indication of the difference between the five major patterns and the other ‘minor’ structures, a summary of the upcoming cases of the paradigmatic variation is presented in Table 4.2. In this chapter, 257 examples will be discussed, all showing one of the five major patterns for the first person complex. Of the minor patterns of the first person complex, I know of only 8 incidental cases, including every curious reference that I have come across. These 8 cases almost all have a different paradigmatic structure, adding to the impression that they are only incidental cases (these incidental cases were discussed in section 3.6.6).

	<i>Major patterns</i> <i>1st pers. compl.</i>	<i>Minor patterns</i> <i>1st pers. compl.</i>	<i>Total</i>
<i>Number of examples</i>	257	8	265
<i>Paradigmatic structures</i>	55	6	61

Table 4.2: Frequencies of the major and minor patterns of the first person complex

The five major patterns of the first person complex can be combined in various ways with the remaining group categories (2+3 and 3+3) and the singular categories (1, 2 and 3). In total, I have found 55 different paradigmatic structures with these eight categories. The 55 different paradigmatic structures are not equally distributed over the world’s linguistic variation. I will make a twofold distinction as to how frequent a pattern is. First, a certain pattern can be ‘common’. There are two criteria for a pattern to be common: the pattern should be attested widely dispersed over the world’s languages and it should be a typical pattern for at least a few genetic families. The first criterion assures that the pattern is found independently in different languages, beyond the influence of common genetic origin or areal contact. The second criterion assures that the pattern is relatively stable, as the paradigms attested are not all exceptional variants within their genetic group. Of the 55 different paradigmatic structures that will be discussed in this chapter, nine will be classified as ‘common’. In contrast, a structure is ‘rare’ if there is one, or maybe a few, examples and if these examples are clearly incidental. Typically, in such ‘rare’ cases, genetically closely related languages do not show the same paradigmatic structure.

The distinction between ‘common’ and ‘rare’ is not rigorously clear cut. The judgments that I present in this chapter are to be interpreted as a proposal to bring some order in the bewildering number of possibilities. The two classes are part of a continuum, which is only divided into these discrete classes to make it more accessible. The continuous nature of the variability makes the boundary between ‘common’ and ‘rare’ at times rather troublesome. There are a few ‘semi-common’ cases that fall in between ‘common’ and ‘rare’. These ‘semi-common’ paradigmatic structures do not occur often, but still there is a sense in which they are more common than the other ‘rare’ structures. Judging from the frequency, these ‘semi-common’ paradigms are nearer to the class of ‘rare’ paradigms than to the class of ‘common’ paradigms. Likewise, the boundary between ‘rare’ and ‘non-attested’ is troublesome. Although many different paradigms are attested, there are still many more theoretically possible structures that are not attested. I will not give any interpretation to the fact that some structures are

attested, albeit rarely, and others are not attested. The circumstance that a particular ‘rare’ paradigm occurs in the present sample, as opposed to other paradigms that do not occur, is probably incidental. If I had studied more languages, or if there had been more languages in the world, some more ‘rare’ cases would probably turn up. When confronted with the wide variety of linguistic structures present in the world’s languages, I do not see any reason why a certain pattern would be impossible for a language. There are clearly structures that are *improbable*, but nothing seems to be *impossible*. I will present all rare patterns that have been attested to show the inherent variability of linguistic structure. As a result, all generalisations that will be formulated in the next chapter will have some counterexamples. Typological generalisations are never to be interpreted as ‘universals’ in the strict sense, but as preferred patterns of linguistic structure.

The presentation of all the cases will be ordered along the lines of the marking of the first person complex. There are five different sections, each devoted to one of the five patterns from Table 4.1. Within these five sections, the type of vertical homophony will be the principal guideline for the classification of the various paradigmatic structures. First the ‘split’ patterns will be discussed, ie the patterns that do not have any vertical homophony. Then the paradigmatic structures with a vertical homophony will be presented. The paradigmatic structures that are classified as ‘common’ all turn out to be ‘split’. Separate subsections are devoted to show that each of these ‘common’ paradigmatic structures is indeed frequent among the world’s languages. For these paradigms, the list of paradigms attested is not complete. Given more time and energy, the lists of the common paradigmatic structures could be expanded with more examples. However, at present the main goal is to argue that these patterns are common, not to establish the precise fraction of the world’s languages that have such common paradigmatic structures. Following the discussion of the common patterns, the examples of the various ‘rare’ and ‘semi-common’ paradigmatic structures are presented. These listings of rare patterns are intended to show the inherent variability of the structure of human language.

4.3 Variants of the ‘unified-we’ type

4.3.1 Preamble

The ‘unified-we’ type of the first person complex is characterised by combined marking of all referential values for ‘we’. The three different kinds of ‘we’ – 1+2, 1+2+3 and 1+3 – are all marked with the same specialised morpheme. The English pronoun ‘we’ is an example of such a unified form.

$1+2$	A	–	A	A	A
$1+2+3$			–	B	C
$1+3$			–	B	C

Table 4.3: The ‘unified-we’ pattern

All possible combinations of the ‘unified-we’ with the categories 2+3 and 3+3 are shown in Table 4.4. There are five possible distributions, all of which are attested among the languages of the world. The first pattern is called ‘split’ as it distinguishes three non-singular categories. The other four patterns are called ‘homophonous’ patterns as some of the three non-singular categories are marked by one morpheme.

	<i>split</i>	<i>homophonous patterns</i>			
		<i>1/2</i>	<i>1/3</i>	<i>2/3</i>	<i>unified</i>
<i>1+2</i>	A	A	A	A	A
<i>1+2+3</i>					
<i>1+3</i>			B		
<i>2+3</i>	B			B	
<i>3+3</i>	C	B	A		

Table 4.4: Non-singular marking patterns for ‘unified we’

These five patterns do not exhaust the possible variation of a pronominal paradigm. More specifically, these five patterns only classify the possibilities of VERTICAL homophony. A pronominal paradigm can also show SINGULAR homophony, HORIZONTAL homophony, or both. All the different kinds of homophony will figure prominently in the discussion of the various examples. The discussion will be streamlined by following the form of the VERTICAL homophony, as this will turn out to be an important factor for the cross-linguistic variability of paradigmatic structure.

First, the SPLIT paradigms will be discussed. Among these paradigms, there are three paradigmatic structures that have a special typological status, for two reasons. First, they occur relatively often and widespread throughout the world’s languages. The examples come from various families from different parts of the globe, a distribution that can only be explained by assuming that the patterns arose multiple times independently of each other. Second, the examples are found to be characteristic for the languages throughout a few linguistic families, indicating that the patterns are not all incidental cases. Each of the three common variants is characterised by one particular paradigmatic structure. They will be called the LATIN-TYPE paradigm, the SINHALESE-TYPE paradigm and the BERIK-TYPE paradigm. These paradigmatic structures are schematically presented in Figure 4.4. Not only are these three pattern widespread over the world’s languages, even more important is that all cases show exactly the same paradigmatic structure. The number of cases is high and the amount of variation is low. This distinguishes these three patterns from the other patterns to be discussed later. In the discussion of these structures, a rough outline of the geographical and genetic stratification of the three common patterns will be presented. The references only indicate that these paradigmatic structures are indeed commonly found among the languages of the world and give an impression how the occurrence of each pattern is distributed over the world. It is rather easy to find examples of these patterns, and it will probably be fairly easy to expand the list of examples that are presented. The Latin-type paradigm will be discussed in section 4.3.2, the Sinhalese-type paradigms

in section 4.3.3, and the Berik-type paradigm in section 4.3.4. Other variants of split paradigms will be presented in section 4.3.5.

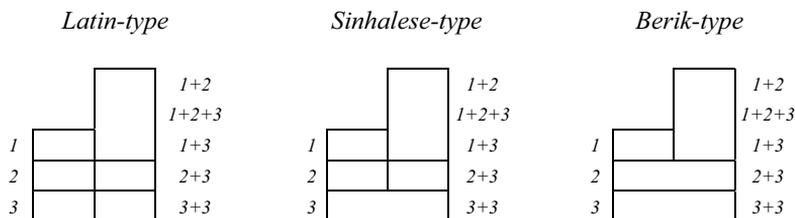


Figure 4.4: Common paradigmatic structures with 'unified we'

The various patterns with a VERTICAL HOMOPHONY will be discussed in sections 4.3.6 to 4.3.9. These patterns are clearly less common than the 'split' pattern. Some of the paradigmatic structures are still found widely distributed over the globe and they might even be attested as a general characteristic of a particular linguistic family, but the number of examples per paradigmatic structure is low. A few of these cases are characterised as 'semi-common'. However, most of these 'vertical' homophonous paradigms are only scarcely found; the examples are mostly atypical within their genetic family and show many different kinds of singular and horizontal homophony. In general, these patterns are not unheard of, but they are clearly not a common thing to happen to a human language. Section 4.3.10 summarises the presented variation of the 'unified-we' paradigms.

4.3.2 Latin-type paradigm

The Latin-type paradigm is well known as a structure of pronominal paradigms. In these paradigms, there is a threefold distinction between the non-singular categories 'we', 'you-all' and 'they'. These categories are distinct from each other and they are also distinct from the singular marking. A prototypical example of this pattern is the Latin inflection. In (4.1), the suffixes of the Latin present indicative are shown. The same inflectional paradigmatic structure can still be found in the inflection of the southwestern Romance languages, as in Portuguese, Catalan and Italian (Posner, 1996:39-43).⁴ The Latin-type of paradigmatic structure is often considered the prototypical pattern of all pronominal marking. As I will show in the rest of this chapter, it is in fact only one of the patterns that can be found regularly among the languages of the world. Other patterns can also be regarded as typical of the structure of human language. The Latin-type paradigm is indeed a common pattern, but it is not the only common one.

⁴ Many paradigm of independent pronouns in West European language have roughly this kind of paradigm. However, there are almost always gender distinctions in the third person singular and sometimes even in other categories, like gender in the Spanish first and second person plural (Plank & Schellinger, 1997:70). In some cases, there are even homophonous forms between singular and plural categories, like 'you' in English or 'sie/Sie' in German.

(4.1) LATIN

1	...-o	...-mus	$I+2$
2	...-s	...-tis	$I+2+3$
3	...-t	...-unt	$I+3$
			$2+3$
			$3+3$

Paradigms of the Latin -found widespread among the Indo-European inflections, although they are not so widespread as is often assumed.⁵ The Indo-European independent pronouns often have gender distinctions, or use demonstrative elements in the third person. Such paradigms are close to the Latin-type, but they are not prototypical examples. Still on the Eurasian continent, paradigms of the exact Latin-type are common in the Uralic inflection.⁶ It is also found in the Turkish inflection (Lewis, 1967: 106-107). Most African languages have gender distinctions (or class marking) in their pronominal paradigms. Due to this extensive paradigmatic variation in the third person, most African paradigms are not included in this chapter. Cases of paradigms without gender or class marking can be found among the independent pronouns of the Mande languages and the Nilotic languages.⁷ Many Southeast Asian languages do not have grammaticalised morphemes for group marking, and consequently these do not have a Latin-type paradigm. Modern Chinese developed a six-way independent person system, and in some variants there is even an inclusive-exclusive distinction (Norman, 1988: 117-121, 157-158). Further south, into the Pacific, there are only exceptional examples of a Latin-type paradigm among the Austronesian languages, because most pronominal paradigms have an inclusive/exclusive opposition.⁸ In contrast, paradigms of the Latin-type seem to be widespread among the Papuan languages from New Guinea.⁹ Finally, Latin-type paradigms are found throughout the American languages, but they are not as common in this area as elsewhere in the world.¹⁰

⁵ Eg among the Iranian languages (Sims-Williams, 1998:144-149), among the Slavic languages (Andersen, 1998:445-446) and in the Brythonic languages Breton (Stephens, 1993:373-375) and Early Cornish (George, 1993:446-447).

⁶ Short descriptions of Uralic languages can be found in Abondolo (1998b). Most Uralic languages have a Latin type inflection. See, for example, Estonian (pp. 140-141), Finnish (pp. 171-174), Mordva (pp. 197-201), Mari (pp. 229-231) and Udmurt (pp. 289-291).

⁷ For Mande, eg Vai (Welmers, 1976:43), Mauka (Ebermann, 1986:74), Koranko (Kanstenkolz, 1987:17-20) and Kpelle (Westermann, 1924:14-15). For Nilotic, eg Maasai (Tucker & Mpaayei, 1953:200), Lotuho (Muratori, 1938:72), Southern Nilotic (Rottland, 1982:137-139, 151, 194, 229, 248-249) and Dinka (Nebel, 1948:15).

⁸ Lynch (1998:101) mentions the independent pronouns of the Siau family in the West Sepik Province of Papua New Guinea (eg Sera, Sissano, Ali, Tumleo and Ulau-Suain), Kiribati and possibly one or two varieties of Fijian.

⁹ A few random examples of Latin-type paradigms from new Guinea are the Daga inflection and independent pronouns (Murane, 1974:34,42-44,63-69), the Waskia future inflection (Ross & Paol, 1978:68) and the Asmat subject suffix (Voorhoeve, 1965:85). This type of pronominal paradigm seems to be widespread, as can be seen from the surveys of Papuan languages by (Voorhoeve, 1975) and (Wurm, 1975b).

¹⁰ Eg the West Greenlandic intransitive inflection (Fortescue, 1984:288), the Ika independent pronouns (Frank, 1990:26) and in the Epena Pedee independent pronouns (Harms, 1994:186,58).

4.3.3 Sinhalese-type paradigm

The next common type of paradigmatic structure is only slightly different from the Latin-type paradigm from the previous section. The defining characteristic of the Sinhalese-type paradigm is that it does not have a specialised element for the category 3+3, the ‘third person plural’. There are two specialised forms for non-singular marking: one for the first person complex, and one for the second person plural. The third person plural is marked together with the third person singular. An example of this paradigmatic structure is shown in (4.2). This is the a-conjugation in the Rumanian inflection (Posner, 1996:43).

(4.2) RUMANIAN

			<i>1+2</i>
		...-m	<i>1+2+3</i>
<i>1</i>	...-Ø		<i>1+3</i>
<i>2</i>	...-i	...-tsi	<i>2+3</i>
<i>3</i>	...-a		<i>3+3</i>

In the case of independent pronouns, it is possible (but not necessary) for the categories 3 and 3+3 to be marked outside the pronominal system. Both categories can be marked by demonstratives, or other linguistic material, which strictly speaking does not belong to the pronominal domain. An example of this type of paradigm is shown in (4.3). These are the independent pronouns from Sinhalese, an Indo-Aryan language from Sri Lanka (Gair, 1970:32-33). I will consider these two, the inflectional and the independent version in (4.2) and (4.3) to be instances of the same paradigmatic structure. I call this type of paradigm the SINHALESE-TYPE paradigm.

(4.3) SINHALESE

			<i>1+2</i>
		api	<i>1+2+3</i>
<i>1</i>	mamə		<i>1+3</i>
<i>2</i>	ohee	oheela	<i>2+3</i>
<i>3</i>	(demonstratives)		<i>3+3</i>

The Sinhalese-type paradigm is found in Eurasia, albeit scarcely. Next to Rumanian and Sinhalese, some more cases are attested among Indo-Aryan languages where third persons are demonstratives (Masica, 1991:251). It is also found in some other languages on the Eurasian continent that do not have third person marking within the pronominal domain, like, for example, the independent pronouns from Buriat, an Altaic language from Siberia (Poppe, 1960:49-52) or Lezgian, a Nakh-Dagestanian from the Caucasus (Haspelmath, 1993). Sinhalese-type paradigms are also uncommon in Africa. It is, however, a general characteristic of the Southern Nilotic subject prefixes. The structure with five different prefixes is found throughout the Southern Nilotic languages. The third person is always overtly marked in Datooga and Omotik, but only occasionally in the Kalenjin languages; in the latter it is often zero. The reconstructed forms for Proto-Southern Nilotic from (Rottland, 1982:243-244) are shown in (4.4).

(4.4) SOUTHERN NILOTIC

			<i>I</i> +2
		*kɪ/kɛ-...	<i>I</i> +2+3
<i>I</i>	*ɑ-...		<i>I</i> +3
<i>2</i>	*ɪ-...	*ɔ-...	<i>2</i> +3
<i>3</i>	*kɔ/ø-...		<i>3</i> +3

Paradigms of the Sinhalese-type are common among North American languages. It is, for example, generally found in the Salish pronominal systems (Newman, 1980:156) and in the Muskogean pronominal affixes. It is exemplified here in (4.5) with the patient prefixes from the Muskogean languages Chickasaw (Payne, 1982:359).¹¹

(4.5) CHICKASAW

			<i>I</i> +2
		po-...	<i>I</i> +2+3
<i>I</i>	sa-...		<i>I</i> +3
<i>2</i>	chi-...	hachchi-...	<i>2</i> +3
<i>3</i>	∅-...		<i>3</i> +3

A few other examples are added to show the common nature of this paradigmatic structure in America. A Sinhalese-type paradigm is found in the Nootka indicative suffixes (Swadesh, 1936:82), shown here in (4.6). It was probably also found in the extinct South American language Muisca (González de Pérez, 1987:74). The independent pronouns from Muisca are shown in (4.7).

(4.6) NOOTKA

			<i>I</i> +2
		...-(m)in	<i>I</i> +2+3
<i>I</i>	...-(m)ah		<i>I</i> +3
<i>2</i>	...-(m)eʔic	...-(m)eʔico	<i>2</i> +3
<i>3</i>	...-ma		<i>3</i> +3

(4.7) MUISCA

			<i>I</i> +2
		chie	<i>I</i> +2+3
<i>I</i>	hycha		<i>I</i> +3
<i>2</i>	mue	mie	<i>2</i> +3
<i>3</i>	as		<i>3</i> +3

Finally, there are also a few examples of languages from New Guinea with a pronominal paradigm of the Sinhalese-type. This pattern is, for example, found in the independent pronouns from Sentani (Cowan, 1965:16) and Asmat (Voorhoeve, 1965:143). The Asmat independent pronouns are shown in (4.8).

¹¹ The same structure is also found in other Muskogean languages; eg Alabama (Lupardus, 1982:66-74) Koasati (Kimball, 1985:107) and Choctaw (Nicklas, 1974:31).

(4.8) ASMAT

			<i>I</i> +2
		me	<i>I</i> +2+3
<i>I</i>	da		<i>I</i> +3
2	wa	ma	2+3
3	na		3+3

4.3.4 Berik-type paradigm

The third common paradigmatic structure has even more horizontal homophony. In this type of paradigm, only one non-singular morpheme is grammaticalised. A morpheme exists for the first person complex, nothing else. In contrast, the other two group categories (2+3 and 3+3) are marked by the corresponding singular morphemes. This means that the morpheme for 2 also marks the category 2+3 and the morpheme for 3 also marks the category 3+3. This paradigmatic structure will be called a BERIK-TYPE paradigm. This paradigmatic structure is exemplified here by the independent pronouns from Berik, a Tor language spoken on New Guinea. Shown in (4.9) is the paradigm referred to as ‘Subject₁’ in the description. There are many more formally and functionally different forms of the independent pronouns, but they all show this same paradigmatic structure (Westrum & Wiesemann, 1986: 38-39).

(4.9) BERIK

			<i>I</i> +2
		ne	<i>I</i> +2+3
<i>I</i>	ai		<i>I</i> +3
2	aame		2+3
3	je		3+3

The same paradigmatic structure is also found in Kuman, another language from New Guinea. The independent pronouns from Kuman are shown in (4.10), as presented by Foley (1986: 70, citing Piau 1985).¹² The pronouns from Kuman resemble the ones from Berik, but they are at most distantly related and spoken in completely different regions of the Island.

(4.10) KUMAN

			<i>I</i> +2
		no	<i>I</i> +2+3
<i>I</i>	na		<i>I</i> +3
2	ene		2+3
3	ye		3+3

Before I proceed with the other examples of the Berik-type paradigm, a word of caution is in place. The Berik-type paradigm is highly similar to another paradigm that will be discussed later in this chapter, the ‘Sierra Popoluca-type’. Some extra care has

¹² This language is referred to as ‘Kamanagu’ by Capell (1940:54) and Forchheimer (1953:65-66).

to be taken to ensure a paradigm is of the Berik-type. The two similar paradigms are shown schematically in Figure 4.5. They might appear rather different, but in practice they are easily mistaken for each other.

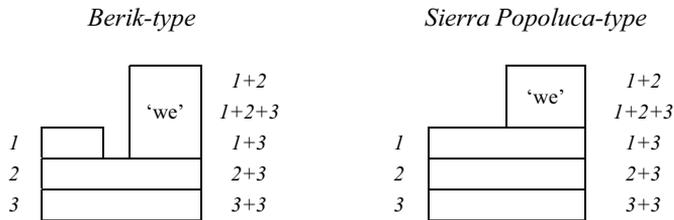
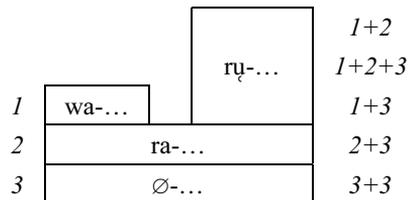


Figure 4.5: The difference between The Berik-type and the Sierra Popoluca-type

The significant correspondence between the two paradigms is that they have only one non-singular morpheme, and in both cases this morpheme is to be translated into English as ‘we’. Taken at face value, both paradigms seem to consist of four different elements: ‘I’, ‘you’, ‘he/she/it’ and ‘we’. However, the two variants of ‘we’ are rather different. In the case of the Berik-type, this morpheme has both the inclusive and the exclusive reading. In the Sierra Popoluca-type, the morpheme ‘we’ has only an inclusive reading. The main argument to classify a paradigm as a case of the Berik-type, and not as a case of the Sierra Popoluca-type, is by transitive uses of this ‘we’, like in sentences as ‘we see you’. An important presupposition of this sentence is that the object ‘you’ is not part of the subject ‘we’. The sentence ‘we see you’ can normally only be interpreted if the ‘we’ has exclusive reading. This fact can be used to distinguish the two paradigms. If the specialised form of ‘we’ can be used in sentences like ‘we see you’, then the ‘we’ can also be used to mark the exclusive. In this case, the paradigm is of the Berik-type.

Paradigms of the Berik-type are found in the Siouan language family from North America. It is exemplified in (4.11) with the active prefixes from Mandan (Mixco, 1997: 17). As can be seen from the transitive sentence (4.12), the meaning of the prefix *ry-*... can also be exclusive. The addressee is the object of the verb ‘to tell’, so this addressee is probably not part of the subject ‘we’.

(4.11) MANDAN



(4.12) MANDAN

ry-ri-kiraktoʔš

EXCL-2-tell

‘We will tell you’

(Mixco, 1997: 17)

This structure is also found in another Siouan language, Assiniboine. The active prefixes from Assiniboine are shown in (4.13), taken from Levin (1964: 31-32). The pre-

fix $y(k)$ -... can have both inclusive and exclusive reading, as shown in (4.14). Example (4.14 a) is explicitly described as having an inclusive meaning. Example (4.14 b) is a transitive sentence with an addressee as object. Because the addressee is in object function here, it is not part of the subject. This means that the ‘we’ subject does not include the addressee; the ‘we’ subject is exclusive in reference. The same structure as in Assiniboine is also described for Lakhota by Van Valin (1977:5, 10-13). Although not all Siouan languages belong to this type, still there are numerous examples to be found in this family.

(4.13) ASSINIBOINE

1	wa-...	y(k)-...	I+2
			I+2+3
2	ma-...		I+3
3	∅-...		2+3

(4.14) ASSINIBOINE

a. *yk-ákxa*

IPLUR-chop

‘I and you chop off something’

(Levin, 1964:31)

b. *cazé-y-ni-yata-pi*

?-IPLUR-2-call-PLUR

‘We call you’

(Levin, 1964:64)

Another group of languages with a pronominal paradigm of the Berik-type is the Eastern Nilotic family, a subgroup from the Nilo-Saharan stock in Africa. The structure of the subject prefixes from these languages is exemplified here with the prefixes from Turkana (Dimmendaal, 1982:120), shown in (4.15). From the example sentences in (4.16), it can be seen that the prefix *ki*-... can have both an inclusive and an exclusive reading. There is an inclusive-exclusive difference in the independent pronouns, and *ki*-... can be used co-referentially with both.¹³

(4.15) TURKANA

1	a-...	ki-...	I+2
			I+2+3
2	i-...		I+3
3	e`-...		2+3

(4.16) TURKANA

a. *kì-losì* *suà*

IPLUR-go EXCL

‘We (exclusive) will go’

¹³ For other Eastern-Nilotic languages, the same structure is found in Teso (Tucker & Bryan, 1966:470), Maasai (Tucker & Mpaayei, 1953:53) and Lotuho (Tucker & Bryan, 1966:470). For a general discussion of the form of the prefixes, see Dimmendaal (1983:279).

- b. *kì-losì* *ηwðñì*
 I PLUR-go INCL
 ‘We (inclusive) will go’
- (Dimmendaal, 1982: 122)

Finally, I know of one incidental case of a Berik-type paradigm from South America. Shown in (4.17) are the ‘person markers’ from Xokleng, a Gé language from Brazil (Urban, 1985: 167).¹⁴

(4.17) XOKLENG

1	nũ	nã	1+2
2	mã		1+2+3
3	wũ		1+3
			2+3
			3+3

It is not so easy to find examples of the Berik-type, compared to the abundance of examples of the previously discussed Latin-type and Sinhalese-type. Still, there are two families where the Berik-type is found throughout the family, and there are some incidental cases. There is a good geographical spread of these examples over the world. This is enough to grant this type the status of common, although it should be kept in mind that it is clearly less common than the two previous types.¹⁵

4.3.5 Rare variants of split ‘unified-we’

There are a few examples of the split-type that do not belong to one of the three common paradigmatic structures. These leftovers are all incidental patterns. They will be discussed in turn.

English-type paradigm

Probably the most famous case is the paradigm of the English pronouns, shown in (4.18). The formerly plural element *ye/you* has taken over the reference of the singular *thou/thee* (Howe, 1996: 170-175). Exactly the same structure is attested in South American Spanish pronouns, where the formerly plural pronoun *vos* has taken over the reference of the singular *tu*. The same change did take place in the history of Dutch, but the distinction between a singular and plural second person has been restored by the addition of a new plural pronoun *jullie* (Howe, 1996: 220-227). Because these well known and extensively described languages all three underwent this

¹⁴ These are probably the same forms as the ‘nominative’ forms as described by Wiesemann (1986b:363). However, the paradigm as presented by Wiesemann also distinguishes number in the second and third person, and gender in the third person, so it might be that Xokleng is not a good example of a Berik-type paradigm. I decided to include it here as Xokleng is at least a candidate for a Berik-type paradigm, and should be considered in this context.

¹⁵ Forchheimer (1953:65-66) also mentions the Korean independent pronouns as an example of the Berik-type paradigmatic structure. Based on more recent descriptions, the independent pronouns from Korean are not a good example of a Berik-type paradigm, although they might historically be derived from such a structure. Synchronically, the independent pronouns have a separate stem for the first person plural in non-deferential contexts. The other non-singular forms do not have a separate stem. However, these non-singulars are obligatorily marked by a number suffix (Sohn, 1994:284-287). I interpret this number suffix as part of the pronominal paradigm. Consequently, the paradigm is not an example of the Berik-type, but of the Latin-type (see also section 3.3).

change, it is sometimes thought off as a common change. It might be a common change, but its end-state is not common at all. The paradigmatic structure of English pronoun, as shown in (4.18), is not attested commonly. Even for English, the overarching form ‘you’ is mainly found in the standard language. In a thorough study of present-day English by Wales (1996), it is noted that ‘many dialects and varieties of English seem to find a formal distinction extremely useful between a single addressee and more than one, at least for statements and questions’ (Wales, 1996: 16-17, 73).¹⁶

(4.18) ENGLISH

		we	<i>I</i> +2 <i>I</i> +2+3
<i>I</i>	I		<i>I</i> +3
<i>2</i>	you		<i>2</i> +3
<i>3</i>	he/she/it	they	<i>3</i> +3

Another case of the same paradigmatic structure as in English is also found in the ‘pronouns’ from Xokleng, a Gé language from Brazil. Note that just as in English, there is a gender distinction in the third person singular (Urban, 1985: 167).¹⁷

(4.19) XOKLENG

		ãŋ	<i>I</i> +2 <i>I</i> +2+3
<i>I</i>	ẽŋ		<i>I</i> +3
<i>2</i>	a		<i>2</i> +3
<i>3</i>	tí/ði	ɔŋ	<i>3</i> +3

Vanimo-type paradigm

A somewhat odd-looking variant is shown in (4.20). These are the actor prefixes from Vanimo, a Sko language from Papua New Guinea.

(4.20) VANIMO

		n-...	<i>I</i> +2 <i>I</i> +2+3
<i>I</i>	∅-...		<i>I</i> +3
<i>2</i>	m-...	∅-...	<i>2</i> +3
<i>3</i>	h/b-...	d-...	<i>3</i> +3

¹⁶ The ‘English-type’ structure is a possibility for human language, but it is not a commonly attested possibility. Compare the structure of the ‘English-type’ to the opposite structures of the ‘Ainu-type’ that will be discussed in section 4.4.2. The Ainu-type paradigm has no opposition in the first and the third person; there is only a difference between singular and non-singular in the second person. In the present sample, the Ainu-type kind of paradigm is roughly just as often found as the English-type paradigm.

¹⁷ In the description of Wiesemann (1986b:363), these pronouns are called ‘non-nominative’. It may seem like there is a separate 2+3 form in this description, but that form is made by adding the plural marker *mē*. This plural morpheme *mē* is an independent element in most Gé languages, and is strictly speaking not part of the pronominal paradigm (Popjes & Popjes, 1986: 185; Wiesemann, 1986b: 361).

In this case, there is a homophony between the second person non-singular and the first person singular. This homophony also is an extremely uncommon structure; I know of no case besides Vanimo. Note that there is a masculine-feminine distinction in the third person singular (Foley, 1986: 134, citing Ross 1980).

Spanish-type paradigm

Finally, there are a few incidental examples that have a ‘split’ non-singular, but a homophony in the singular. This type of paradigm is, for example, found in the Spanish ‘imperfecto’ suffixes, presented in (4.21). Roughly half of the different suffixal paradigms in Spanish have this structure. The other half has a Latin-type structure with three different singular forms (Hallebeek *et al.*, 1994: 114). Exactly the same paradigmatic structure, with a first-third person homophony in the singular, but three different forms in the non-singular is found in the Icelandic ‘weak preterite’ (Thráinsson, 1994: 159).

(4.21) SPANISH

1	...-ba	...-bamos	$1+2$
2	...-bas	...-bais	$1+2+3$
3	...-ba	...-ban	$1+3$
			$2+3$
			$3+3$

Icelandic-type paradigm

Still another singular structure is found in the Icelandic ‘weak present’, as shown in (4.22). In this case, there is a different homophony in the singular compared to the Spanish case above (Thráinsson, 1994: 159).

(4.22) ICELANDIC

1	...-∅	...-dum	$1+2$
2	...-ur	...-duð	$1+2+3$
3		...-du	$1+3$
			$2+3$
			$3+3$

Siciliano-type paradigm

The last example that I want to present is the paradigm of the ‘imperfect’ suffixes from Siciliano, San Fratello dialect, shown in (4.23). This case is interesting as it shows that it is possible for a singular homophony and a horizontal homophony to appear without a vertical homophony in the non-singular. However, it remains the only case of such a structure that I know off (Bigalke, 1997: 60).

(4.23) SICILIANO

1	...-va	...-mu	$1+2$
2	...-vi	...-vi	$1+2+3$
3	...-va	...-vu	$1+3$
			$2+3$
			$3+3$

These last three paradigmatic structures seem to be extremely unusual. If there are three different specialised forms of the non-singular, then there is a strong tendency for there to be also three different forms in the singular. Horizontal homophony is common among split paradigms, but singular homophony is only exceptionally found in some European languages.

4.3.6 The 1/2-homophony of ‘unified-we’

After the discussion of the split paradigms, I will now turn to the vertical homophonous variants of ‘unified we’. The different kinds of vertical homophony are repeated in Table 4.5. These four cases will be discussed in turn. Among these paradigms, singular homophony is relatively common but horizontal homophony is extremely uncommon. Six paradigmatic structures are slightly less rare compared to the others. These six will be categorised as ‘semi-common’.

	<i>1/2</i>	<i>1/3</i>	<i>2/3</i>	<i>unified</i>
<i>1+2</i>	A	A	A	A
<i>1+2+3</i>				
<i>1+3</i>		B	B	
<i>2+3</i>		A		
<i>3+3</i>	B			

Table 4.5: Vertical homophonous patterns of ‘unified we’

The first kind of homophony to be discussed is characterised by a combination of the first person complex with the second person plural into one specialised morpheme. There is one genetic family where this pattern is found consistently. Other than that, I have found seven incidental examples of this homophony, distributed widely dispersed over the world. This is not as frequent as the previous three ‘common’ patterns, but still common enough to take note of. These cases are distributed over five different paradigmatic structures. The paradigms with a 1/2-homophony show a large amount of variation. Within this variation, the most widely occurring structure has three separate singular morphemes, a structure presented in Figure 4.6 as the SLAVE-TYPE paradigm. Among the cases presented in this section, there are a few paradigms that also have a homophony of the singular categories. All logically possible singular homophonies are attested, but the most widely occurring singular homophony is the mirror-image of the non-singular. This structure is presented in Figure 4.6 as the SVAN-TYPE paradigm.

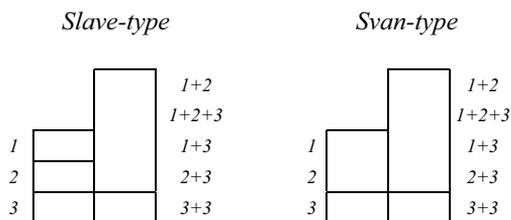


Figure 4.6: The semi-common structures with 1/2-homophony

Slave-type paradigm

Paradigms of the Slave-type are attested as the regular structure of the independent pronouns of the Athabascan languages. The pronouns from the Slavey variant of Slave are presented in (4.24). The same paradigmatic structure can be found in many Athabascan languages.¹⁸ The prefixal object marking in all Athabascan languages shows the same structure, throughout the family. However, these prefixal paradigms distinguish two different kinds of third person marking – regular and ‘obviative’ (eg Rice, 1989:253,431).

(4.24) SLAVE

1	s _ɨ	nax _ɨ	1+2
2	n _ɨ		1+2+3
3	ʔed _ɨ		1+3
		ʔeged _ɨ	2+3
			3+3

I know of two examples outside the Athabascan family where exactly the same Slave-type paradigmatic structure is found. The first is shown in (4.25). These are the independent pronouns from Awa, a highland language from Papua New Guinea. (Loving, 1973:85).

(4.25) AWA

1	ne	ite	1+2
2	ade		1+2+3
3	we		1+3
		se	2+3
			3+3

The last example of the Slave-type is the pronominal inflection from the Tommo-so variant of Dogon, shown in (4.26). There are many different portmanteau forms for the category 3+3, fused with aspect and an eventual negation. The presented form ...-eŋ is the ‘positive stative’ morpheme (Plungian, 1995:30).

(4.26) DOGON

1	...-m	...-y	1+2
2	...-w		1+2+3
3	...-∅		1+3
		...-eŋ	2+3
			3+3

Svan-type paradigm

In the following paradigmatic types, the singular categories show some kind of homophony. These examples of singular homophony are all found in inflectional paradigms; a generalisation that was already established in the previous chapter. The most

¹⁸ Other examples of the Slave-type among the Athabascan languages are found in the independent pronouns of Chiricahua Apache (Hoijer, 1946:76,78,83), Navaho (Young & Morgan, 1987:7-8), Kato (Goddard, 1912:33) and Hupa (Goddard, 1905:29).

obvious case is a homophony between the first and second person singular. This singular homophony is a mirror image of the non-singular homophony. This structure is exemplified with the imperfect suffixes from the South Caucasian language Svan, shown in (4.27). Note that there are also pronominal prefixes in Svan that co-occur with these suffixes. These prefixes are of a completely different paradigmatic type, to be discussed in section 4.5.2 (Tuite, 1997:28).

(4.27) SVAN

1	...-äs	...-ad	1+2
2			1+2+3
3	...-a	...-ax	1+3
			2+3
			3+3

There are two more cases with the same Svan-type structure. The first case is found in the Papuan language Waskia, shown in (4.28). The suffixes presented are the past simple suffixes (Ross & Paol, 1978:67-68).

(4.28) WASKIA

1	...-em	...-man	1+2
2			1+2+3
3	...-am	...-un	1+3
			2+3
			3+3

The final example of a Svan-type paradigm is found in the North American language Nez Perce. In the perfective and irrealis, the number prefix *pe-...* is found together with the pronominal prefix *hi-...*. The resulting paradigm is shown in (4.29) (Rude, 1985:30-39).

(4.29) NEZ PERCE

1	∅-...	pe-...	1+2
2			1+2+3
3	hi-...	hipe-...	1+3
			2+3
			3+3

Lak-type paradigm

A slight variation on the Svan-type is found in Lak, a Nakh-Dagestanian language from the Caucasus. The non-past suffixes from Lak are shown in (4.30). A horizontal homophony is added here, this time between the third person singular and non-singular (Helmbrecht, 1996:131).

(4.30) LAK

1	...-ra	...-ru	$1+2$
2			$1+2+3$
3	...-ri		$1+3$
			$2+3$
			$3+3$

Dongola-type paradigm

Not all examples with a singular homophony show a mirror-structure between the singular and the non-singular. There are also some cases with a vertical 1/2-homophony that have a different homophony in the singular. The first of these is found in Kenuzi-Dongola, an Nilo-Saharan language from Sudan. The suffixes shown in (4.31) mark the present of the Dongola dialect. This paradigm shows a different homophony in the singular; the forms for 2 and 3 are identical (Reinisch, 1879:65).

(4.31) KENUZI-DONGOLA

1	...-ri	...-ru	$1+2$
2	...-in		$1+2+3$
3		...-ran	$1+3$
			$2+3$
			$3+3$

Ika-type paradigm

The last example of a 1/2-homophony comes from South America. The subject prefixes from Ika, a Chibchan language from Peru, are shown in (4.32). In this case, there is a homophony in the singular forms between 1 and 3 (Frank, 1990:50-51).

(4.32) IKA

1	∅-...	a-...	$1+2$
2	nʌ-...		$1+2+3$
3	∅-...	ri/win-...	$1+3$
			$2+3$
			$3+3$

To conclude, there turns out to be rather a lot of examples of the 1/2-homophony, but these examples show a great deal of internal variation. Only the Slave-type and the Svan-type paradigms have been attested in more than one instance. However, the frequency of these two paradigmatic structures comes nowhere near the frequency of the split-paradigms. The Slave and Svan paradigms are classified as ‘semi-common’.

4.3.7 The 1/3-homophony of ‘unified-we’

Characteristic for the 1/3-homophony is the similarity between the first person complex and the third person non-singular. This is not really a common pattern; it is only regularly found in a group of languages on the southeastern tip of Papua New Guinea. Outside that area, I know of only a few languages that show this homophony. The two paradigmatic structures in Figure 4.7 are the most common among this restricted set of examples. The OMIË-TYPE paradigm distinguishes three different morphemes for

the singular categories. The MIDDLE DUTCH-TYPE paradigm shows the 1/3-homophony analogously in the singular.

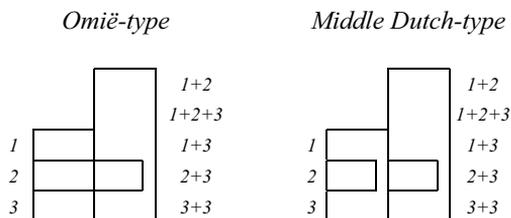
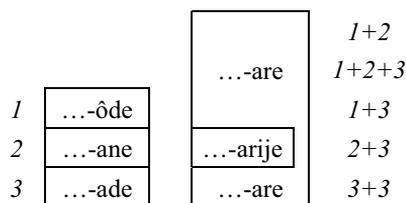


Figure 4.7: The semi-common structures with 1/3-homophony

Omië-type paradigm

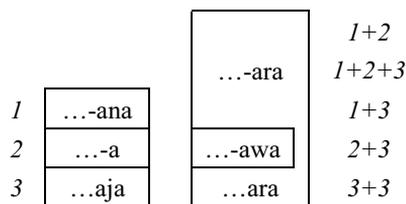
The first of the 1/3-homophony paradigms is exemplified in (4.33) with the past suffixes from Omië, a Papuan language from the southeastern tip of Papuan New Guinea (Austing & Upia, 1975: 544).

(4.33) OMIË



An Omië-type paradigm is also found in the pronominal inflection of the neighbouring language Orokaiva. The Orokaiva marking is exemplified here with the so called ‘indicative Mid Past B’ paradigm, shown in (4.34). Other suffixal paradigms in Orokaiva have the exactly the same paradigmatic structure (Healey *et al.*, 1969: 62). This paradigmatic structure is also found in the related language Korafe (Farr & Farr, 1975: 747-749). Orokaiva and Korafe belong to the Binanderean family, which is only distantly related to Koiarian family to which Omië belongs. However, the resemblance between the paradigms points to at least a strong areal influence.

(4.34) OROKAIVA



A last example of an Omië-type paradigm is found in Bagirmi, a Nilo-Saharan language from Chad. The presented forms in (4.35) are the independent pronouns as found in isolation (Gaden, 1909: 10).¹⁹

¹⁹ Non-singular marking with a 1/3-homophony possibly also exists in inflection of the Toro variant of Dogon, a Niger-Congo language from Mali, although the description is not completely clear on this

(4.35) BAGIRMI

1	ma		d'e	1+2
2	i		se	1+2+3
3	ne		d'e	1+3
				2+3
				3+3

Middle Dutch-type paradigm

In some cases, the homophony between 1 and 3 is mirrored in the singular. This is, for example, found in the inflection of Middle Dutch. Shown in (4.36) are the suffixes of the Middle Dutch 'praeteritum' (Schönfeld, 1959:144-145). The same paradigms, also with singular 1/3 homophony, is attested in the past inflection of contemporary German (Eisenberg, 1994:367-371).

(4.36) MIDDLE DUTCH

1	...-∅		...-en	1+2
2	...-es		...-et	1+2+3
3	...-∅		...-en	1+3
				2+3
				3+3

Another example of a Middle Dutch-type structure is the Omië present, shown in (4.37). The same homophony between 1 and 3 is found both in the singular and in the non-singular (Austing & Upia, 1975:544)

(4.37) OMIË

1	...-aje		...-arue	1+2
2	...-anue		...-aruje	1+2+3
3	...-aje		...-arue	1+3
				2+3
				3+3

Bagirmi-type paradigm

Finally, there are a few cases of a 1/3-homophony in which there is horizontal homophony as well. In these cases, the non-singular category 2+3 is combined with a singular category. Such a link between singular and non-singular is attested in two different version. The first variant is paradigm of the subject clitic from Bagirmi, shown in (4.38). In this clitic paradigm, the marking of 2+3 is identical to 2; they are both zero. The subject clitic shown here is a reduced version of the pronoun; cf (4.35) above (Gaden, 1909:10).

[continued from previous page]

point (Calame-Griaule, 1968:xxxv). Note that there is a completely different set of inflectional person markers described for the Tommo-so variant of Dogon, discussed as (4.26) above.

(4.38) BAGIRMI

			$1+2$
		d'-...	$1+2+3$
1	m-...		$1+3$
2	∅-...		$2+3$
3	n-...	d'-...	$3+3$

German-type paradigm

Another link between the singular and the non-singular is attested in German. The present suffixes from German are shown in (4.39). Different from the Bagirmi structure above, in German the categories 2+3 and 3 are marked identically (Eisenberg, 1994: 367-371). In the Sudanic language Midob, the suffixes for complete action show exactly the same structure as the German suffixes in (4.39), with a homophony between 2+3 and 3 (Thelwall, 1983: 107). This paradigm is not shown here.

(4.39) GERMAN

			$1+2$
		...-en	$1+2+3$
1	...-e		$1+3$
2	...-st	...-t	$2+3$
3	...-t	...-en	$3+3$

4.3.8 The 2/3-homophony of 'unified-we'

Characteristic of this type is the conflation of the categories 2+3 and 3+3, resulting in an opposition between 'we' and 'non-we'. The non-singular 2/3 homophony is not that widely attested. It is rather often documented for highland Papuan languages of New Guinea. However, in most of these cases, also dual forms are included in the paradigm. The discussion of these paradigms with dual forms is postponed to section 8.3.3. Outside New Guinea, I know of only a few incidental cases, mainly to be found in America. The most widely occurring structure with a 2/3-homophony is a paradigm of the NEZ PERCE-TYPE, as shown in Figure 4.8. In another 'semi-common' set of the examples that will be discussed below, the structure of the singular marking mirrors the non-singular. Such paradigms have an KOMBAL-TYPE structure with a homophony in the singular marking, conflating the singular categories 2 and 3.

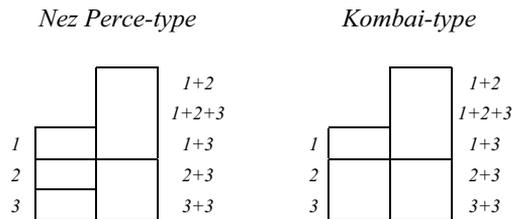


Figure 4.8: The semi-common structures with 2/3-homophony

Nez Perce-type paradigm

The first case of this homophony is the ‘unmarked’ form of the Nez Perce independent pronouns, shown in (4.40). There are other paradigms of independent pronouns in Nez Perce, besides the one presented here, but all have exactly the same paradigmatic structure (Rude, 1985:123).

(4.40) NEZ PERCE

1	'íin	núun	1+2
2	'íim	nún	1+2+3
3	'ipí	'imé	1+3
			2+3
			3+3

Exactly the same Nez Perce-type paradigm is found in Warekena, an Arawakan language from Venezuela. The independent pronouns from Warekena are shown in (4.41), but the same structure is also found in the inflectional marking. Note that the third person singular is marked by demonstrative elements. This pronominal pattern is probably an innovation of Warekena, as it is not found in any other Arawakan language (Aikhenvald, 1998:293,322).

(4.41) WAREKENA

1	nuya	waya	1+2
2	piya	waya	1+2+3
3	(demon.)	niya	1+3
			2+3
			3+3

It has been claimed in the literature that this pattern is typical for languages from the highlands of New Guinea. For example, Foley (1986:72) calls it ‘the typical highlands conflation of second and third persons in the nonsingular’. He even reconstructs a paradigm of independent pronouns for Proto-Gorokan with a Nez Perce-type paradigm, although it should be noted that, purely synchronically spoken, none of the Gorokan languages has this kind of independent pronouns (Foley, 1986:248-249). Indeed, in the verbal inflections of the highland languages of New Guinea, a homophony between 2+3 and 3+3 is often found, but almost always in combination with a dual with the same homophony.²⁰

Kombai-type paradigm

The only clear example without dual marking of a conflation between 2+3 and 3+3 that I have found in New Guinea is the suffixal paradigm from Kombai, as shown in (4.42). The morphemes can probably be analysed as a first person marker ...-f and a plural marker ...-o (de Vries, 1989:145). Note that this paradigm is different from the

²⁰ Eg the present neutral from Yagaría (Renck, 1975:87-88), the independent pronouns from Wiru (Foley, 1986:72, citing Kerr 1966) and the subject suffixes from Fore (Foley, 1986:74, citing Scott 1978); see also Wurm (1975b:476). These paradigms will be discussed in the chapter on dual marking (see especially section 8.3.3).

Nez Perce paradigm in (4.40) because there is a homophony in the singular mirroring the non-singular.

(4.42) KOMBAL

1	...-f	...-fo	1+2
2	...-∅	...-o	1+2+3
3			1+3
			2+3
			3+3

The Kombai-type paradigm is also found in the pronominal inflection of Chitimacha, an extinct language from the USA, shown in (4.43). Just as with the suffixes from Kombai above, the suffixes from Chitimacha may be analysed as consisting of a separate number ...-*nV*- and person morpheme ...-*Vk* (Swadesh, 1946:317-318).

(4.43) CHITIMACHA

1	...-ik	...-nuk	1+2
2	...-i	...-na	1+2+3
3			1+3
			2+3
			3+3

The last examples of this type comes from South America. In (4.44) the pronominal prefixes from Lengua, a Mascoian language from Paraguay, are shown. Note that there is considerable morphophonological variation; the presented prefixes are only one of the possible series of forms. There does not seem to be a possibility to separate person from number marking in this case (Susnik, 1977:98).

(4.44) LENGUA

1	ik-...	nīn-...	1+2
2	ap-...	kyel-...	1+2+3
3			1+3
			2+3
			3+3

Koryak-type paradigm

Finally, there is one special case of a paradigm with a 2/3-homophony. This example is found in the Chukotko-Kamchatkan languages, here exemplified in (4.45) by the indicative intransitive prefixes from Koryak (Comrie, 1980a: 64,67). Instead of separate singular and non-singular forms, both are marked identically by zero in Koryak.

(4.45) KORYAK

1	t-...	mæt-...	1+2
2	∅-...		1+2+3
3			1+3
			2+3
			3+3

4.3.9 The unified-homophony

The final kind of vertical homophony is the unified-homophony. Characteristically, paradigms with a unified-homophony have only one specialised morpheme for all non-singular categories. In a sense, this is a combined form of all the previous kinds of vertical homophony. This structure is rather rare and highly varied. The few examples that I know of all have a different marking structure of the singular categories.

Lango-type paradigm

The first example of a structure with a unified homophony is the paradigm of the habitual prefixes of Lango, a Nilotic language from Sudan, shown in (4.46). Comparative data indicate that this paradigm has lost some of its non-singular distinctions relatively recently (Bavin, 1981:90-91).

(4.46) LANGO

1	à-...	ò-...	1+2
2	ì-...		1+2+3
3	∅-...		1+3
			2+3
			3+3

Dutch-type paradigm

The other cases of a unified-homophony that have been attested show also a homophony in the singular. A paradigm with a singular homophony between 2 and 3 is found in the pronominal inflection of Standard Dutch. The suffixes for the present, without inversion, are shown in (4.47). The usage of a second person non-singular suffix *...-t* (cf German in (4.39) above) can still be found incidentally, but is considered to be rather old fashioned.

(4.47) DUTCH

1	...-∅	...-en	1+2
2	...-t		1+2+3
3	...-t		1+3
			2+3
			3+3

Waskia-type paradigm

Almost the same structure is found in the present suffixes of the Papuan language Waskia, shown in (4.48). Other than in the case of Dutch, there is a singular homophony between 1 and 2 in Waskia (Ross & Paol, 1978: 67-68).

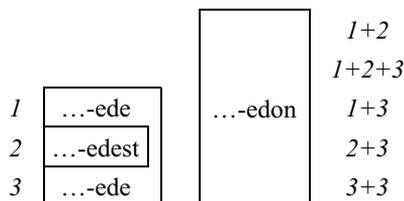
(4.48) WASKIA

1	...-sam	...-san	1+2
2	...-so		1+2+3
3	...-so		1+3
			2+3
			3+3

Old English-type paradigm

Finally, the Old English past suffixes show a singular homophony between 1 and 3 (Robertson & Cassidy, 1954: 141). Concluding, the unified-homophony, that is characterised by one morpheme for all non-singular categories, is attested throughout the world, but with a wide variety of singular paradigmatic structures.

(4.49) OLD ENGLISH

**4.3.10 Summary**

To conclude this wide variety of examples, this section will summarise the major typological features of the ‘unified-we’ paradigms. The frequencies that are reported in this section are counts of the presented examples. In Appendix B, a detailed account of these numbers is presented.

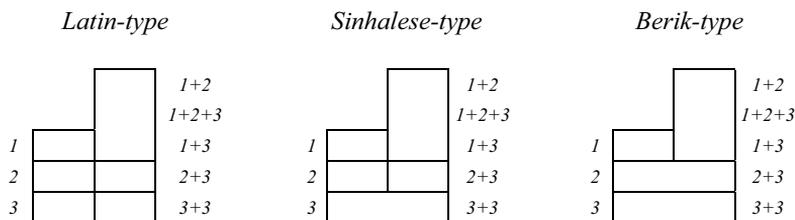


Figure 4.9: common paradigmatic structures with ‘unified-we’

In Figure 4.9, the three common paradigmatic structures with ‘unified-we’ are repeated. These three structures are attested frequently throughout the world’s languages. Structurally, the LATIN-TYPE can be interpreted as the most ‘complete’ type, relative to which the SINHALESE-TYPE loses the distinction between singular and non-singular in the third person. The BERIK-TYPE paradigms also neutralises the singular/non-singular difference in the second person. This structural hierarchy is mirrored in the attested frequency of these three patterns. As can be seen in Table 4.6, there is a difference in frequency between the three common paradigms. The Latin-type is more frequent than the Sinhalese-type, which is on its turn more frequent than the Berik-type.

	<i>number of cases</i>
<i>Latin-type</i>	38
<i>Sinhalese-type</i>	19
<i>Berik-type</i>	10
Total	67

Table 4.6: Number of cases of the common paradigms

The special status of the three common paradigms relative to all other paradigmatic structures is striking. The three common types account for 67 examples. This means they occur on an average in 22.3 cases per paradigmatic structure. All other attested paradigmatic structures together account for 41 examples. These 41 examples are distributed over 20 different paradigmatic structures. This is roughly ten times less frequent than the common paradigmatic structures. However, the uncommon paradigmatic structures are not all equally uncommon. The distinction between common and uncommon is more like a continuum than a clear cut division. In Figure 4.10 below, the most frequent of the ‘rare’ paradigmatic structures are presented. These six structures account for 22 of the uncommon cases, a mean of 3.7 examples per structure (the remaining 15 uncommon structures account for 19 cases, a mean of 1.3 examples per structure). The amounts for the three classes are summarised in Table 4.7.

	<i>common</i>	<i>semi-common</i>	<i>rare</i>
<i>examples</i>	67	22	19
<i>paradigmatic structures</i>	3	6	15
<i>examples per structure</i>	22.3	3.7	1.3

Table 4.7: Frequencies of unified ‘we’ paradigms

Generalising on the ‘semi-common’ structures, the notable characteristic is that they do not show any horizontal homophony. In all six structures, the singular and non-singular are strictly separated. In contrast, all these structures show vertical homophony. Either the first person complex is homophonous with 2+3 (the Slave-type), or the first person complex is homophonous with 3+3 (the Omië-type), or 2+3 and 3+3 are homophonous (the Nez Perce-type). In the other cases, the vertical homophony is mirrored in the singular.

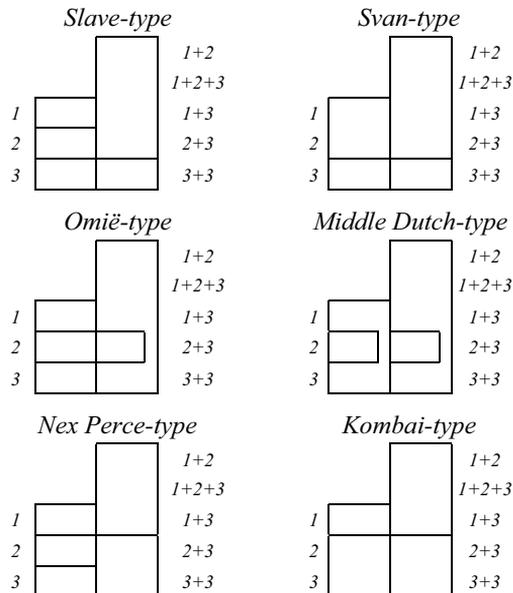


Figure 4.10: Semi-common paradigmatic structures with ‘unified-we’

4.4 Variants of the ‘no-we’ type

4.4.1 Preamble

The ‘no-we’ type is characterised by non-existing marking for the first person complex. It is rather unusual (though not impossible) for a whole language not to mark any form of ‘we’ (see section 3.6.2). However, for a single paradigm within a language, no specialised form for ‘we’ is frequently attested. The variability of these paradigms will be laid out in this section.

<i>I+2</i>			A	A	A
<i>I+2+3</i>	A	–			B
<i>I+3</i>			–	B	C

Table 4.8: The ‘no-we’ type

When the ‘no-we’ first person complex is combined with the categories 2+3 and 3+3, there are five theoretical possibilities, as shown in Table 4.9. The dashes represent categories that are not marked with a specialised non-singular morpheme. All first person complexes are marked with a dash in the table, this being the definitional property of the present class of paradigms. The capital letters in the table represent categories that are marked differently from the first person complex.

	<i>split</i>	<i>homophonous patterns</i>			
		<i>1/2</i>	<i>1/3</i>	<i>2/3</i>	<i>unified</i>
<i>I+2</i>					
<i>I+2+3</i>	–	–	–	–	
<i>I+3</i>					–
<i>2+3</i>	A		A		
<i>3+3</i>	B	A	–	A	

Table 4.9: Theoretical possible marking patterns of ‘no-we’

The first of these theoretical possible pattern is called ‘split’ as the three non-singular categories are marked differently. The other four patterns show some kind of homophony between the three categories. These patterns are indicated by the kind of homophony, for example ‘1/2’ meaning that the first person complex and the second person non-singular are marked identically. Only four of the five possibilities from Table 4.9 are attested among the world’s linguistic variation. I have not been able to find an example of the 1/3-homophony. The other homophonous patterns are also relatively rare, so the absence of even a single example of the 1/3-homophony is probably only due to chance. The split-type is by far the most common pattern. This variation of the various split paradigms will be discussed in section 4.4.2. The few examples of homophonous examples will be presented in section 4.4.3. The variation of the all these ‘no-we’ paradigms will be summarised in section 4.4.4.

Salt-Yui-type paradigm

A split ‘no-we’ paradigm is also frequently attested among independent pronouns. In these paradigms, it is often difficult to say whether the marking of the group-categories are taken over by the singular pronouns, or whether there is no marking at all for the group-categories (cf section 2.6.4 on the third person singular). Both options seem to occur, although a clear dividing line between these options can not be drawn. The independent pronouns of the American language Acoma Keresan are a case where the extension into the non-singular pronouns does not seem possible. The Acoma pronouns only distinguish two forms: *hínun*^{ʔé} and *hišum*^{ʔé}. These two words are strictly used to mark speaker and addressee. They can not be used for ‘we’ or ‘you-plural’ respectively. This pronominal paradigm only consists of two elements, the rest of the categories are non-existing (Maring, 1967:43-44, 113-114). A different, and more common, situation is found in Salt-Yui, a non-Austronesian language from Papua new Guinea. Just as in Acoma, there are only two independent pronouns. In Salt-Yui, the two pronouns are *na* and *ni*. These pronouns are used for speaker and addressee reference. There is no third person pronoun; other linguistic material is used to mark this function. Also, there is no number distinction in this pronominal paradigm.

‘Two personal pronouns occur distinguishing first and second person. Third person is shown by the specific noun optionally followed by a demonstrative, that is, there is not third person pronoun as such. ... Number is not distinguished in the pronouns.’ (Irwin, 1974:32)

However, in the case of Salt-Yui, the independent pronouns can also be used for plural reference. As shown in (4.52a), the first person pronoun can also be used for ‘we’. In (4.52b), a second person pronoun is used for ‘you-plural’. A plural noun *yasu*, meaning ‘two people’, is added for clarification. In another example (not shown here) the noun *yalhobi*, meaning ‘fellows’ is used to clarify the plurality of *ni* (Irwin, 1974:74). The paradigmatic structure of the independent pronouns of Salt-Yui, as shown in (4.51), is probably identical to the Maricopa structure, as shown in (4.50). Some caution has to be taken, as there is no clear example of an inclusive use of the pronoun *na* in the grammar.

(4.51) SALT-YUI

1	na	1+2
2	ni	1+2+3
3	(demonstratives)	1+3
	(?)	2+3

[continued from previous page]

Pame, an Oto-Manguean language from Mexico (Manrique, 1967:343-344). In South America, the inflectional prefixes from Ayoreo, a Zamucuan language from Paraguay, are of this type (Susnik, 1973:52-57). Finally, the pronominal marking of Mura Pirahã is of this type. Interestingly, in Pirahã there is no number marking whatsoever (see section 3.6.2).

(4.52) SALT-YUI

a. *na ama ha holo wamga ...*
 1 girls talk collecting walk
 ‘When we court the girls ...’ (Irwin, 1974:135)

b. *ni yasu ala mol dibilge*
 2 two people inside COP say
 ‘You are both inside!’ (Irwin, 1974:61)

In many languages, number is not marked inside the pronominal paradigm. Often plurality is only marked as an overt addendum to pronouns. In Salt-Yui, words like ‘people’ and ‘fellows’ are used to mark plurality. In Golin, a language from New Guinea related to Salt-Yui, the word *kobe*, meaning ‘people’ is used (Foley, 1986:70, citing Bunn 1974).²³ In many Southeast Asian languages, there are so-called associative markers that can be used to produce group-marking reference from singular referential elements (see section 3.3). Such associative markers turn a pronoun ‘I’ into something like ‘I and my fellows’, which can be translated as ‘we’, although only in the exclusive meaning. In modern Cantonese a suffix *...-dei* is used, as exemplified in (4.53). This suffix is not only found with independent pronouns, but also with proper names, like *Ling* in the example below. The same situation is found in Japanese, using a suffix *...-tachi*.²⁴

(4.53) CANTONESE

ngo I
ngo-dei I and associates, ‘we’
lei you
lei-dei you and associates, ‘you-all’
ah Ling dei Ling and his associates, Ling and that crowd

In Vietnamese, the nominal element *dowq* meaning ‘group’ is used to mark plurality of independent pronouns. In Vietnamese, this addition seems to be obligatory to mark plural reference (Cooke, 1968:76-77). In the case of Vietnamese, the associative marking is grammaticalised as part of the pronominal paradigm. Consequently, it is not a case of ‘no-we’ anymore, but an example of a Latin-type paradigm. In other languages, however, the addition to disambiguate nouns is only needed when the context requires it. The optional use of associative marking is also found in Classical Chinese, where a word meaning ‘group’ or ‘associates’ could be used to overtly mark plurality (Norman, 1988:89-90). A final example of independent pronouns without overt marking of plurality is Thai. In Thai, many pronouns can be used indifferently as to number. However, there are a few pronouns that, when used in a specific honorific setting, are basically singular or plural.

²³ Separate marking for plurality of pronouns is also found in Manem, another non-Austronesian language from Papua New Guinea, not directly related to Salt-Yui and Golin (Voorhoeve, 1975:416). The short grammatical notes in the source do not mention the precise function of this pluraliser.

²⁴ These data on associative marking come from a data file compiled by Edith Moravcsik, collected through a query on the Linguist list, see Moravcsik (1994). On Cantonese, see also Norman (1988:219-220). On Japanese, see also Shibatani (1990:371-372).

‘Personal pronouns, like nouns, do not have number as an obligatory formal category, but not all forms are ambiguous as to number ... [eg the form] /raw’/ (first person) ... occurs freely in either singular or plural usage, with or without overt plural expressions accompanying them. However, in singular usage, /raw’/ is a somewhat restricted first person form used chiefly by royalty speaking to commoners, or between intimate equals. In plural usage it is a general term which may serve as the corresponding plural of all first person singular forms except terms used to royalty.’ (Cooke, 1968:19)

A split in independent pronouns is a rather widespread phenomenon. The preceding examples show that is common in Southeast Asia and in New Guinea. It is also regularly found among the languages of America.²⁵ The independent case of ‘split’ together with the inflectional cases from the Maricopa-type form a common type of paradigmatic structure.

Big Nambas-type paradigm

Finally, there are a few examples of split-type paradigms that have neither a Maricopa-type paradigm, nor a Salt-Yui-type paradigm. These patterns are incidental cases. There is one case that distinguishes singular from non-singular only in the third person. Shown in (4.54) are the person prefixes from Big Nambas, a language spoken on the island Malekula, part of Vanuatu in the Pacific. There are other affixes in the language that indicate number, so the difference between the singular and the plural is marked regularly, though not in the pronominal paradigm as such (Fox, 1976:52-61).

(4.54) BIG NAMBAS

				<i>1+2</i>
				<i>1+2+3</i>
<i>1</i>		n-...		<i>1+3</i>
<i>2</i>		kə-...		<i>2+3</i>
<i>3</i>	i-...		a-...	<i>3+3</i>

Ainu-type paradigm

There are a few cases that only distinguish singular from non-singular in the second person. The first example comes from Classical Ainu, shown in (4.55). The presented prefixes are used for transitive subject marking (Shibatani, 1990:25). Contemporary Ainu has developed an inclusive-exclusive opposition.

(4.55) CLASSICAL AINU

				<i>1+2</i>
				<i>1+2+3</i>
<i>1</i>		a-...		<i>1+3</i>
<i>2</i>	e-...		eci-...	<i>2+3</i>
<i>3</i>	∅-...			<i>3+3</i>

²⁵ Southeast Asian and New Guinean examples have been discussed extensively in the preceding section. Let me add two more examples from South America to stretch the geographical distribution of this kind of pronominal paradigm. The independent pronouns from Xerente, A Gé language from Brazil, is probably of the Salt-Yui type, although the description is not completely clear on this point (Wiesemann, 1986b:361,364-368). The independent pronouns from Paez, a language from Colombia, is clearly described as a Salt-Yui type paradigm, although there are different forms for gender (Rojas Curieux, 1991:52-53).

Exactly the same paradigmatic structure is found in the independent pronouns of Tairora, a highland language from New Guinea, shown in (4.56). These pronouns are related to the pronouns from Usarufa, as shown below in (4.60). The data on Tairora are from Foley (1986:255). Still another example with the same structure is the paradigm of the personal affixes from Gubden, a dialectal variant of Dargi, shown in (4.57). Dargi is a Nakh-Dagestanian language spoken in Dagestan. This paradigm is related to the suffixes of literary Dargi, as shown below in (4.61). The data on Gubden are from Helmbrecht (1996: 138).

(4.56) TAIRORA

				$1+2$
				$1+2+3$
1	te			$1+3$
2	are		be	$2+3$
3	bi			$3+3$

(4.57) GUBDEN

				$1+2$
				$1+2+3$
1	...-ra			$1+3$
2	...-de		...-da	$2+3$
3	...-∅			$3+3$

Suki-type paradigm

A very special case is the paradigm of the independent pronouns from the Papuan language Suki, shown in (4.58). The first person complex is homophonous with the second person singular (Foley, 1986:72, citing Voorhoeve 1970).

(4.58) SUKI

				$1+2$
				$1+2+3$
1	ne		e	$1+3$
2	e		de	$2+3$
3	u		i	$3+3$

Aztec-type paradigm

A slightly different structure is shown in (4.59). These are the pronominal prefixes from Classical Nahuatl, an Aztec language from Mexico. The same ‘diagonal’ homophony as in Suki is found here. The only difference is the homophony in the third person. Note that there is also a suffix ...-? that disambiguates the homophony of ‘you-singular’ and ‘we’ (Newman, 1967:193). This structure is commonly found in the contemporary Aztec languages.²⁶

²⁶ Cf the Milpa Alta variant of Nahuatl as described by Whorf (1946:384) and Pipil, as described by Campbell (1985:54).

(4.59) CLASSICAL NAHUATL

			<i>I</i> +2
		ti-...	<i>I</i> +2+3
<i>I</i>	ni-...		<i>I</i> +3
<i>2</i>	ti-...	am-...	<i>2</i> +3
<i>3</i>	∅-...		<i>3</i> +3

4.4.3 Homophonous variants of ‘no-we’

After the large number of split patterns, I now turn to the few examples of homophonous variants. There are four different theoretical possibilities of vertical homophony, repeated here in Table 4.10. The possibility of a 1/3-homophony is not attested in the current sample. However, the number of examples of each of the attested kinds of homophony is generally low when compared to the relative ubiquity of the common paradigms. The absence of the 1/3-homophony is likely to be a chance phenomenon. In contrast to the homophonous paradigms with ‘unified-we’ (see section 4.3), there are no ‘semi-common’ paradigmatic structures among the following cases.

	<i>1/2</i>	<i>1/3</i>	<i>2/3</i>	<i>unified</i>
<i>I</i> +2				
<i>I</i> +2+3	–	–	–	
<i>I</i> +3				–
<i>2</i> +3		B		
<i>3</i> +3	B	–	B	

Table 4.10: Homophonous patterns of ‘no-we’

Usarufa-type paradigm

A few cases of ‘no-we’ paradigms use the same morpheme for the first person complex and for the second person non-singular. Such a paradigmatic structure is attested in the independent pronouns from Usarufa, a highland language from New Guinea, shown in (4.60); data from Foley (1986:255). This paradigm is related to the independent pronoun from Awa, which was presented above in (4.25).

(4.60) USARUFA

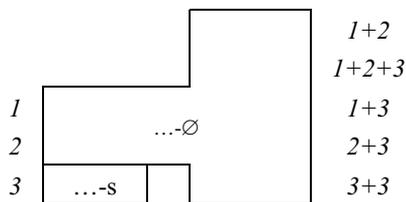
			<i>I</i> +2
			<i>I</i> +2+3
<i>I</i>	ke		<i>I</i> +3
<i>2</i>	e		<i>2</i> +3
<i>3</i>	we	ye	<i>3</i> +3

Dargi-type paradigm

A slightly different pattern is found in literary Dargi. In this case, the third person non-singular is marked together with the third person singular. It is possible to disambiguate the 2+3 form by using a suffix ...-ya, but this suffix is not obligatorily used (van den Berg, 1999: 155-156). Some dialectal variants of Dargi have a separate morpheme for 2+3; for example Gubden, shown above in (4.57).

found in English. The present inflection of English verbs is shown in (4.65). All group-categories are marked identical, by zero, and this same zero is also found in the singular.

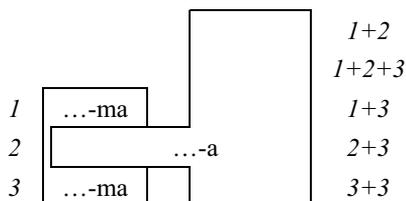
(4.65) ENGLISH



Koiari-type paradigm

The unified-homophony seems to be rare. The only other example that I know of is found in Koiari, a non-Austronesian language from Papua New Guinea. The present/imperfect suffixes from Koiari are shown in (4.66). Just as in English, all group-categories are marked identical, with a suffix *...-a*, and this same suffix is also used for the second person singular. The same paradigmatic structure is also found in the past/perfect, with different morphemes (Dutton, 1996:23).

(4.66) KOIARI



4.4.4 Summary

Two common paradigmatic structures were discussed: the Maricopa-type and the Salt-Yui-type. The only reason to separate these two structures was their morphological status. The Maricopa-type paradigms are inflectional; the Salt-Yui-type paradigms are morphologically independent. However, structurally these two types are identical. From now on, I will consider both types to be variants of one paradigmatic structure as shown in Figure 4.12.

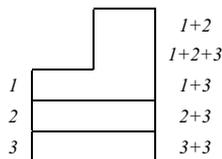


Figure 4.12: The Maricopa/Salt-Yui-type paradigmatic structure

Many other paradigmatic structures are attested, but these are almost all incidental cases. The examples that are presented in the text are summarised in Table 4.11. The reported frequencies are accounted for in Appendix B. Twenty cases are presented of the common Maricopa/Salt-Yui-type paradigmatic structure. Other than that, 16 examples were discussed of 11 different uncommon paradigmatic structures, a mean of 1.5 examples per structure. The frequencies of the different varieties of 'no-we' para-

digms are comparable to the variation as was found with ‘unified-we’ paradigms in the previous section (cf Table 4.7). The common paradigms have roughly 20 examples per structure, the rare cases have roughly 1.5 examples per structure. The main difference is that there are no ‘semi-common’ structures attested here. The distinction between common and uncommon is rather clear-cut in the case of ‘no-we’.

	<i>common</i>	<i>rare</i>
<i>examples</i>	20	16
<i>paradigmatic structures</i>	1	11
<i>examples per structure</i>	20.0	1.5

Table 4.11: Frequencies of ‘no-we’ paradigms

4.5 Variants of the ‘only-inclusive’ type

4.5.1 Preamble

The next type of marking for the first person complex is characterised by a special form for the inclusive ‘we’ and the absence of a specialised morpheme for the exclusive counterpart. This is the first kind of paradigmatic structure that distinguishes between inclusive and exclusive. This distinction between inclusive and exclusive will turn out to be a strong predictor for the variability of paradigmatic structure.

<i>I+2</i>			A	A	A
<i>I+2+3</i>	A	–	A	A	B
<i>I+3</i>			–	B	C

Table 4.12: The ‘only-inclusive’ type

The ‘only-inclusive’ marking of the first person complex is attested in one common paradigmatic structure, mnemonically called the SIERRA POPOLUCA-TYPE paradigm. The form of this common paradigmatic structure is presented in Figure 4.13. There are a few examples of other paradigmatic structures, but these are all exceptional cases.

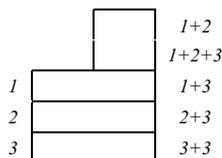


Figure 4.13: Common paradigmatic structure of ‘only-inclusive’

4.5.2 Sierra Popoluca-type paradigm

The only common variant of ‘only-inclusive’ is a pattern with only one specialised morpheme for group-marking. Besides an inclusive ‘we’, there are no other non-

singular morphemes. One pattern is found in the majority of all cases. In this common pattern, the category 1+3 aligns with 1, 2+3 with 2 and 3+3 with 3. This is exemplified in (4.67) with the prefixes from Sierra Popoluca, a Mixe-Zoque language from Mexico (Foster & Foster, 1948:17-19; Elson, 1960:207). By using a plural suffix *...-táʔm*, the difference between the singular reference and the group reference is disambiguated. This is shown in (4.68a, b). Note that it is also possible to use the plural suffix with the inclusive prefix *ta-....* By using this combination, it is possible to distinguish between 1+2 and 1+2+3, as is shown in (4.68c, d).²⁷

(4.67) SIERRA POPOLUCA

		ta-...	1+2
			1+2+3
1	ʔa-...		1+3
2	mi-...		2+3
3	∅-...		3+3

(4.68) SIERRA POPOLUCA

a. *ʔa-moŋ-táʔm-pa*
1-sleep-PLUR-IMPERF
'we (exclusive) are sleeping'

b. *mi-móŋ-táʔm-pa*
2-sleep-PLUR-IMPERF
'you (plural) sleep'

c. *ta-moŋ-pa*
1+2-sleep-IMPERF
'we (inclusive) will sleep'

d. *ta-moŋ-táʔm-pa*
1+2-sleep-PLUR-IMPERF
'we (inclusive plural) will sleep'

(Foster & Foster, 1948:19)

Example sentences like (4.68a) are crucial to classify a pronominal paradigm as 'only-inclusive'. The Sierra Popoluca-pattern is strongly alike to the pattern of the Berik-type. Both patterns have only one non-singular morpheme, which is in both cases to be translated as 'we'. However, the possible referents for this 'we'-morpheme in Sierra Popoluca is different from Berik, as shown in Figure 4.14. The difference between the two patterns is the marking for the category 1+3. In the Berik-type, the category 1+3 is marked identical to the categories 1+2 and 1+2+3, adding up to a general form like the English 'we'. Alternatively, in the Sierra Popoluca-type, the category 1+3 is marked identical to the speaker, leaving 1+2 and 1+2+3 as sole

²⁷ The pronominal system of Sierra Popoluca is described by Foster & Foster (1948:17-19) as being of the minimal/augmented type. In fact, this description is the first known description of a minimal/augmented pronominal system (see section 3.6.5). I do not follow the Fosters in this analysis. By the classification as used in the present work, the pronominal prefixes of Sierra Popoluca are separated from the other affixes (cf Elson, 1960:207). Notably, the separate number suffix is not included in the discussion of the pronominal prefixes. Only when the pronominal prefixes would be considered together with this number suffix, eight different referential values would be distinguished (comparable to the eight different forms of a minimal/augmented paradigm).

meanings of the form ‘we’. Taken at face value, the two types might look alike, as both only have one non-singular form to be translated into English using the pronoun ‘we’. However, the structures are significantly different, and should be clearly distinguished (cf the discussion on page 110 ff.).

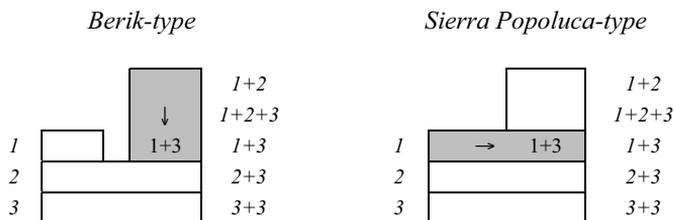


Figure 4.14: The Sierra Popoluca-type compared to the Berik-type

From the graphic form of the Sierra-Popoluca paradigm, it is understandable that these paradigms are often analysed as 4-person systems. Analysed as such, the paradigm from Sierra Popoluca would look as shown in (4.69). This ‘4-person’ analysis is entirely analogous to my analysis, as presented in (4.67), only stripped down to the essentials. I will stick to the more burdensome analysis as in (4.67) for it to be comparable with the other patterns that are discussed in this chapter. However, for the description of individual languages that have a pronominal paradigm of this sort, an analysis like in (4.69) is to be preferred.

(4.69) SIERRA POPOLUCA

$1+2$	ta-...
1	?a-...
2	mi-...
3	∅-...

This paradigmatic structure is a typical pattern of languages of South America. The pronominal prefixes from Trio, a Carib language from Surinam, are shown in (4.70). This is a clear example of a Sierra Popoluca-type paradigm (Carlin, 1997). The same paradigmatic structure is also found in another Carib language from Surinam, confusingly called ‘Carib’, but also known as Kalihna (Hoff, 1968: 134).

(4.70) TRIO

		k-...	$1+2$
			$1+2+3$
1	w-...		$1+3$
2	m-...		$2+3$
3	n-...		$3+3$

A few more examples from South America follow. First, the pronominal inflection of Maká, a Guaicuran language from Paraguay, show this structure (Gerzenstein, 1994: 83-97). Next, the independent pronouns and the pronominal inflection from Jaqaru, an Aymaran language from Peru, both have a Sierra Popoluca-type paradigm (Hardman, 1966: 79). The pronominal prefixes from Canela-Kraho, a Gé language

from Brazil are another example (Popjes & Popjes, 1986:175). Finally, the pronominal affixes and the independent pronouns from the Campa languages (a subgroup of Arawakan) have this paradigmatic structure, although they distinguish two gender forms in the third person (Wise, 1971:67; Payne, 1981:34; Reed & Payne, 1986:324-327). The only example in North America is shown in (4.71): the prefix paradigm from Winnebago, a Siouan language (Greenberg, 1988, citing Susman 1943).

(4.71) WINNEBAGO

		hĩ-...	<i>I+2</i>
			<i>I+2+3</i>
<i>1</i>		ha-...	<i>I+3</i>
<i>2</i>		ra-...	<i>2+3</i>
<i>3</i>		∅-...	<i>3+3</i>

Paradigms of the Sierra-Popoluca type are also found among the Papuan languages of New Guinea. Shown in (4.72) are the pronominal prefixes from Nimboran. The same roots are also used for morphologically independent pronouns (Anceaux, 1965:167). Other examples of this paradigmatic structure in New Guinea are found in the Border family. Both the independent pronouns of Imonda (Seiler, 1985:44) and the short form of the pronouns in Amanab (Minch, 1991:31-32) have a Sierra Popoluca-type structure.

(4.72) NIMBORAN

		io-...	<i>I+2</i>
			<i>I+2+3</i>
<i>1</i>		ŋo-...	<i>I+3</i>
<i>2</i>		ko-...	<i>2+3</i>
<i>3</i>		no-...	<i>3+3</i>

A final example of this type is the paradigm of the subject prefixes from the South Caucasian language Svan, shown in (4.73). Just as in Sierra Popoluca, there are suffixes in Svan to disambiguate the number-reference of these prefixes. These suffixes form an interesting paradigm in themselves; they have been discussed on page 118 (Tuite, 1997:23).

(4.73) SVAN

		l-...	<i>I+2</i>
			<i>I+2+3</i>
<i>1</i>		xw-...	<i>I+3</i>
<i>2</i>		x-...	<i>2+3</i>
<i>3</i>		∅/l-...	<i>3+3</i>

The number of examples of the Sierra Popoluca-type is not overwhelming, but there are enough cases, widely dispersed over the world's languages, to grant this pattern the classification 'common'. This is different from what has been claimed in the literature, where this type is often depicted as rather exotic and exceptional, as, for example, would appear from the comment by Plank (1985). He says that:

‘... diese Paradigmen in einer Hinsicht, das Verhältnis von Person und Numerus betreffend, eher untypisch sind.’ (Plank, 1985:143)

Paradigms of the Sierra Popoluca-type are not atypical structures for pronominal paradigm; on the contrary, they are completely normal patterns for a human language.

4.5.3 Rare variants of ‘only-inclusive’

There is a small group of examples of other paradigmatic structures with an ‘only-inclusive’ first person complex. This group of exceptional examples is not only small, it also consists of incidental examples of each structure. The scarcity and variability of these rare variants clearly illustrates that there is a large difference between the one common paradigmatic structure, the Sierra Popoluca-type, and the other rare variants.

Warrwa-type paradigms

In the actor clitics/prefixes of the Australian language Warrwa exceptional kinds of horizontal homophony are attested. The paradigms for the clitics differ slightly according to the tense and the verb class. In the future of verb class I, a pattern with only a horizontal homophony in the first person is found, shown in (4.74a). In the non-future of verb class I, a horizontal homophony of first and second person is found, shown in (4.74b). Finally, in (4.74c), the future of verb class 2, shows a homophony in the first and third person (McGregor, 1994:41).

(4.74) WARRWA

A.

		ya-...	<i>I</i> +2
			<i>I</i> +2+3
<i>I</i>	nga-...		<i>I</i> +3
<i>2</i>	mi-...	ku-...	<i>2</i> +3
<i>3</i>	∅-...	ngi-...	<i>3</i> +3

B.

		ya-...	<i>I</i> +2
			<i>I</i> +2+3
<i>I</i>	ka-...		<i>I</i> +3
<i>2</i>	wa-...		<i>2</i> +3
<i>3</i>	∅-...	ku-...	<i>3</i> +3

C.

		ya-...	<i>I</i> +2
			<i>I</i> +2+3
<i>I</i>	ka-...		<i>I</i> +3
<i>2</i>	nga-...	wa-...	<i>2</i> +3
<i>3</i>	ku-...		<i>3</i> +3

Algonquian-type paradigm

The remaining examples of rare paradigmatic structures all show some kind of vertical homophony. However, if there is any vertical homophony, then there is also a horizontal homophony. Together, these different kinds of homophony result in intricately structured paradigms. The most famous example of this type is found in the

languages of the Algonquian family. Verbs have pronominal prefixes in the Algonquian languages. The prefixes for intransitive verbs are exemplified in (4.75 a) with data from Southwestern Ojibwa (Schwartz & Dunnigan, 1986:305).²⁸ In this paradigm, the exclusive marking (1+3) is identical to the marking of the speaker (1) and the inclusive (1+2, 1+2+3) is marked differently. In the case of Ojibwa, the inclusive marking (1+2, 1+2+3) is identical to the marking of the second person plural (2+3), which is in turn identical to the singular addressee (2); they are all marked by the prefix *kit-....* It should be added that the Algonquian languages use obligatory suffixes to mark number. In combination with the prefixes, these suffixes disambiguate seven categories. The combined prefixes and suffixes are shown in (4.75 b).

(4.75) SOUTHWESTERN OJIBWA

A.

		kit-...	I+2
			I+2+3
I	int-...		I+3
2	kit-...		2+3
3	∅-...		3+3

B.

		kit-...-min	I+2
			I+2+3
I	int-...	int-...-min	I+3
2	kit-...	kit-... m	2+3
3	∅-...	∅-...-wak	3+3

Huave-type paradigm

Another vertical homophonous case of ‘only-inclusive’ is found in the Mexican language Huave. The pronominal prefixes of Huave are shown in (4.76 a).²⁹ Interestingly, the inclusive is not grouped here with the second person, but with the third person. Again, there are number suffixes. Once these number suffixes are taken into account, as shown in (4.76 b), all eight categories are distinguished (Stairs & Hollenbach, 1969:48-53).

(4.76) HUAVE

A.

		a-...	I+2
			I+2+3
I	sa-...		I+3
2	i-...		2+3
3	a-...		3+3

²⁸ Schwartz & Dunnigan (1986) use the name ‘Ojibwe’ in stead of ‘Ojibwa’. For data on other Algonquian languages, eg Eastern Ojibwa (Bloomfield, 1956:44) and Menomini (Bloomfield, 1962:36-40). Bloomfield (1946:97-99) presents the same structure as a general characteristic of Algonquian languages on the basis of a comparison of Fox, Cree, Menomini and Ojibwa.

²⁹ There is massive allomorphy in these pronominal affixes, and the prefixes that are described here are a strongly idealised picture of the present prefixes of verb class A.

B.

		a-...-ar	$I+2$
		a-...-a:c	$I+2+3$
1	sa-...	sa-...-an	$I+3$
2	i-...	i-...-an	$2+3$
3	a-...	a-...-iw	$3+3$

Lenakel-type paradigm

The last rare variant with an ‘only-inclusive’ first person complex comes from Lenakel, an Austronesian language from Vanuatu, shown in (4.77). The vertical homophony between inclusive and third plural that has been attested in Huave is also found here. However, in Lenakel there is no horizontal homophony in the third person (Lynch, 1978:45). Comparable patterns are found in some other languages of the other Tanna family, the family to which Lenakel belongs (Lynch, 1967:46-48).³⁰

(4.77) LENAHEL

		k-...	$I+2$
			$I+2+3$
1	i-...		$I+3$
2	n-...		$2+3$
3	r-...	k-...	$3+3$

4.5.4 Summary

The ‘only-inclusive’ first person complex includes one common paradigmatic pattern. This common paradigm, repeated here in Figure 4.15, is called the Sierra Popoluca-type of paradigmatic structures. This kind of paradigm is found in a wide variety of languages all over the world.

			$I+2$
			$I+2+3$
1			$I+3$
2			$2+3$
3			$3+3$

Figure 4.15: The Sierra Popoluca-paradigm

Besides the common pattern, a few incidental cases of other paradigmatic structures patterns are attested. The counts as presented in Table 4.13 are specified in Appendix B. The distinction between common and rare is rather clear-cut in this case.

	<i>common</i>	<i>rare</i>
<i>examples</i>	15	7
<i>paradigm structures</i>	1	6
<i>examples per structure</i>	15.0	1.2

Table 4.13: Frequencies of ‘only-inclusive’ paradigms

³⁰ Cf the paradigm from Kwamera that is shown on page 261.

4.6 Variants of the ‘inclusive/exclusive’ type

4.6.1 Preamble

The ‘inclusive/exclusive’ type distinguishes both an overtly marked inclusive and an overtly marked exclusive form of ‘we’. Both forms for ‘we’ are marked by a specialised morpheme; a morpheme that is not used to mark singular referents. In such ‘inclusive/exclusive’ paradigms, at least these two different forms for ‘we’ are specialised non-singular morphemes. Most examples, though, have more non-singular forms besides these two. In this section, I will discuss how the other non-singular categories are morphologically marked

$I+2$			A	A	A
$I+2+3$	A	–	A	A	B
$I+3$			–	B	C

Table 4.14: The ‘Inclusive/Exclusive’ type

Three paradigmatic structures of the inclusive/exclusive type are common. These structures are presented in Figure 4.16. These three structures are attested widely scattered through the world’s linguistic diversity. I will discuss these three common patterns first, in separate subsections. The MANDARA-TYPE will be discussed in section 4.6.2 and the TUPI-GUARANI-TYPE will be discussed in section 4.6.3. The KWAKIUTL-TYPE is attested throughout the world’s languages, but it is never found as the prototypical paradigmatic structure of a whole linguistic family. This pattern is classified as semi-common. The known examples are discussed in section 4.6.4. Other paradigmatic structures are attested, but they are very rare. I know of only incidental examples for each of the rare structures. These rare patterns will be discussed in section 4.6.5.

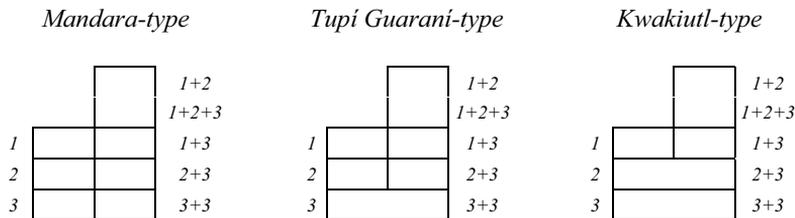


Figure 4.16: (Semi-)common paradigmatic structures with ‘inclusive/exclusive’

4.6.2 Mandara-type paradigm

The paradigmatic structures of the Mandara-type was characterised by Ingram (1978:219) as one of four structures that is ‘more frequent than the others’. Indeed, this paradigmatic structure is found widespread among the world’s languages. Interestingly, the cases with this structure seem to have a clear preference for morphologically independent marking. I know of only a few examples of an inflectional Man-

dara-type paradigms. This paradigm is exemplified in (4.78) with the non-completive pronouns from Mandara, a Chadic language from Cameroon (Burquest, 1986:78).

(4.78) MANDARA

1	yá	má	1+2
		ɲá	1+2+3
2	ká	kwá	1+3
3	á	tá	2+3
			3+3

Patterns of the Mandara-type are indeed numerous and easy to find. In Africa, cases of Mandara-type are attested mainly south of the Sahara. In the Afro-Asiatic stock, it is found among the Chadic and the Cushitic languages.³¹ In the Nilo-Saharan stock, it is described for all languages from the Teso-Turkana subgroup.³² Finally, in the Niger-Congo stock, this type of paradigm is encountered in the basic pronouns of some of the Grassfields languages and in the subject pronouns of some of the Atlantic languages.³³ An inclusive/exclusive opposition might not be a prototypical characteristic of African languages, but still it is quite commonly found (contrary to claims that it is a rare phenomenon in Africa (cf Nichols, 1992:123-124)). In the Eurasian continent, paradigms of the Mandara-type are uncommon, but attested in widely dispersed languages. It is found, for example, in the Caucasus and in Siberia.³⁴ In Chinese, the Mandara-type is encountered in most modern northern varieties, but it was not present in the classical language (Norman, 1988:89,157-158). It is possible that an inclusive developed through areal influence from the Siberian or Mongolian languages. However, the opposite development – loss of the inclusive-exclusive distinction – has occurred in Mongolian (Sanzheyev, 1973:74). The Mandara-type paradigm is not encountered in the Indo-European stock except for independent pronouns in the Indo-Aryan languages Gujarati (Tisdall, 1892:18) and Marathi (Pandharipande, 1997:376,381). Probably, both languages borrowed the inclusive-exclusive opposition from the neighbouring Dravidian languages (Masica, 1991:251). The contact situation between Dravidian and Indo-Aryan shows again (just as in the Chinese-Mongolian situation) that the influence can go both ways. The inclusive/exclusive opposition is not found in the Dravidian language Kannada, probably under influence of the neighbouring Indo-Aryan languages (Sridhar, 1990:205). Interestingly, an inclu-

³¹ For Chadic, eg the independent pronouns from Podoko and Mandara (Burquest, 1986:78,83). For Cushitic, eg the ‘enclitics’ from Somali (Kirk, 1905:29-30,103-104). Note that there is a gender distinction in the Somali third person singular.

³² Eg the independent pronouns from Turkana (Dimmendaal, 1982:207) and Karimojong (Novelli, 1985:107).

³³ For the Grassfields languages, see, for example, Aghem (Hyman, 1979:47,49). This structure is also found in the Beoid language Noni (Hyman, 1981:15). The Mandara-type structure is only found in the basic pronouns of these languages. In both languages, there are also compound pronouns with a rather different structure (see Appendix A). For the Atlantic languages, see, for example, Adamawa Fulfulde (Taylor, 1921:32) and Palor (D’Alton, 1983:201-203).

³⁴ For the Caucasus, eg the independent pronouns from Chechen (Nichols, 1994:32). For Siberia, eg the independent pronouns from the language-isolate Gilyak (Austerlitz, 1959; Gruzdeva, 1998:25-26). A Mandara-type paradigm is also found in the pronominal inflection and in the independent pronouns from the Tungusic languages Even (Malchukov, 1995:12,16) and Evenki (Bulatova & Grenoble, 1999:33).

sive-exclusive opposition is common among the Dravidian languages, but the Mandara-type paradigm is not so often found.³⁵ Most pronominal paradigms of Dravidian languages are of the Tupí Guaraní-type that will be discussed in the next section. Travelling into the Pacific, the Mandara-type paradigm is described regularly for Austronesian languages.³⁶ In contrast to the abundance of cases among the Austronesian languages, it is sometimes assumed in the literature that an inclusive-exclusive distinction is exceptional among the non-Austronesian languages of New Guinea and the surrounding islands (Lynch, 1998:167). However, independent pronouns with a Mandara-type paradigm are attested regularly among the Non-Austronesian languages in the Bird's Head (Vogelkop) of New Guinea, leading over the 'neck' of the Bird's Head towards Mainland, and in Non-Austronesian languages on the islands west of New Guinea.³⁷ This can be explained as an Austronesian influence. However, influence in the opposite direction is also attested. The Siao family, a group of Austronesian languages spoken on the eastern mainland of New Guinea, have lost the inclusive-exclusive distinction, probably under influence of the non-Austronesian languages (Lynch, 1998:101). The Mandara-type paradigm is only incidentally found in Australia, as most pronominal paradigms with an inclusive-exclusive distinction also distinguish dual forms.³⁸ These paradigms with duals will be discussed in section 8.5.2. Finally, the Mandara-type paradigm is described for geographically widely dispersed American languages, although it is not attested very frequently.³⁹ Note that the inclusive/exclusive opposition in any of its various forms is found very often among American languages. It is only this particular Mandara-type paradigmatic structure that turns out to be rather uncommon.

³⁵ A Mandara-type paradigms is found in the pronominal inflection in Koya (Tyler, 1969:92) and in Tamil (Asher, 1982:173-174). Note that both inflectional paradigms have gender distinctions in the third person.

³⁶ Among the Oceanic branch of Austronesian, the Mandara-type paradigm is, for example, found in the independent pronouns from Motu, Mono-Alu, Nakanamanga, Pulawat (Lynch, 1998:100-101) and Big Nambas (Fox, 1976:30). Examples of independent pronouns of this type from Austronesian languages outside the Oceanic branch are Toba Batak (Nababan, 1981:77) and Chamorro (Topping, 1973:107). There are also some Austronesian languages that have an inflectional Mandara-type paradigm eg Uma (Esser, 1964:36), Trukese (Lynch, 1998:103-104), Tondano (Sneddon, 1975:104,113) and Toba Batak (Nababan, 1981:77).

³⁷ Voorhoeve (1975) lists many examples of independent pronouns of the Mandara-type. A large part of these data are from preliminary field reports. However inconclusive, these descriptions indicate that this paradigmatic structure is rather common in Irian Jaya (ie the western part of New Guinea). Voorhoeve mentions as examples the Mairasi-Tanah Merah stock (ibid.:424-425), the West Bomberai stock (ibid.:433-434), the South Bird's Head family (ibid.:438-439) and the Inanwatan family (ibid.:440). The Mandara-type is also found among the Island Non-Austronesian languages, west of Irian Jaya (Capell, 1975:693-696). Both these rather preliminary descriptive surveys do not present any information on inflectional paradigms. I know of two non-Austronesian languages in this area with an inflectional Mandara-type paradigm: the language-isolate Warembori (Donahue, 1999:28-30) and the North Halmahera language Galela (Shelden, 1991:162). There are probably more inflectional cases, but it is difficult to find good data.

³⁸ An incidental case is the paradigm of the pronominal inflection from Gunin (McGregor, 1993:44-46).

³⁹ Eg the independent pronouns from Shuswap, a Salish language from Canada (Kuipers, 1974:59), the pronouns from Ojibwe, an Algonquian, language from the USA (Schwartz & Dunnigan, 1986:296) and the pronouns from Retuarã, a Tucanoan language from Colombia (Strom, 1992:34-35). Note that in Retuarã gender distinctions are found in the third person.

4.6.3 Tupí Guaraní-type paradigm

Paradigms of Tupí Guaraní-type are only slightly different from the previous Mandara-type. The only difference is a homophony in the third person. The categories 3 and 3+3 are marked identically. This homophony is not found extremely often when compared to the abundance of examples of the Mandara-type. Still, there are enough cases to grant the Tupí Guaraní-type the status ‘common’. Paradigms of the Tupí Guaraní-type are the general structure that is found in the intransitive prefixes of the Tupí-Guaraní family in South America.⁴⁰ The paradigms are so strongly alike that the paradigmatic structure can even be confidently reconstructed for proto-Tupí-Guaraní (Jensen, 1990:120). The reconstructed active prefixes are shown in (4.79).

(4.79) PROTO-TUPÍ-GUARANÍ

		*ya-...	1+2
		*oro-...	1+2+3
1	*a-...	*pe-...	1+3
2	*ere-...		2+3
3	*o-...		3+3

The independent pronouns of Dravidian languages are often of the Tupí Guaraní-type. Shown in (4.80) are the pronouns from Toda (Emeneau, 1984:95,98). The third person pronouns are not specialised person markers, but general referential elements that are also used as demonstratives. This paradigmatic structure is a general characteristic of Dravidian languages, although it is not found in all Dravidian languages.⁴¹

(4.80) TODA

		om	1+2
		em	1+2+3
1	o.n	nīm	1+3
2	ni.		2+3
3	(demonstratives)		3+3

The same situation as in Dravidian is found in the Tucanoan languages, spoken around the border of Colombia, Brazil and Ecuador. In some Tucanoan languages, the third person pronouns are identical to the demonstrative pronouns, as, for example, in Macuna (Smothermon *et al.*, 1995:35), Siona (Malone, 1988:126) and Koreguaje (Gralow, 1993:15). In other Tucanoan languages, the third person pronouns and the demonstratives are alike, but not identical (Malone, 1988:126).

⁴⁰ Eg Guajajara (Bendor-Samuel, 1972:86-92), Guaraní (Adelaar & Silva López, 1988:31-35), Aché (Susnik, 1961/62:98), Asurini (Harrison, 1971:30) and Mundurukú (Crofts, 1973:69-70,88-91).

⁴¹ Eg Malayalam (Asher & Kumari, 1997:258,266), Tamil (Asher, 1982:143-144,148-149) and Koya (Tyler, 1969:59-60). For a general account of Dravidian pronominal structure, see Caldwell (1856:414-415,420).

(4.81) MACUNA

		bādi	<i>I</i> +2
		gia	<i>I</i> +2+3
<i>I</i>	ji		<i>I</i> +3
<i>2</i>	bī	bīa	<i>2</i> +3
<i>3</i>	(demonstratives)		<i>3</i> +3

Finally, I will present a few incidental cases of the Tupí Guaraní-type paradigm. First, the imperfective pronouns from Ngizim, a Chadic language from Nigeria, are shown in (4.82). This case is interesting because, differently from the previous examples, the Ngizim independent third person pronoun is not a demonstrative item (Burquest, 1986:76). Next, the transitive subject prefixes from Colloquial Ainu are shown in (4.83); data from Shibatani (1990:28). A final case (not shown here) are the pronominal prefixes from Ticuna, a language isolate spoken mainly in Brazil. Note that the Ticuna paradigms distinguish gender in the third person (Anderson, 1966:6-10).

(4.82) NGIZIM

		wàa	<i>I</i> +2
		jàa	<i>I</i> +2+3
<i>I</i>	naa		<i>I</i> +3
<i>2</i>	kaa	kwaa	<i>2</i> +3
<i>3</i>	aa		<i>3</i> +3

(4.83) COLLOQUIAL AINU

		a-...	<i>I</i> +2
		ci-...	<i>I</i> +2+3
<i>I</i>	ku-...		<i>I</i> +3
<i>2</i>	e-...	eci-...	<i>2</i> +3
<i>3</i>	∅-...		<i>3</i> +3

4.6.4 Kwakiutl-type paradigm

The Kwakiutl-type paradigm is a semi-common pattern of the inclusive/exclusive type. This pattern is found well distributed over the world's diversity, but it is never attested as a general characteristic of a close-knit family. As a result, this paradigm is not common – following the criteria for commonness that have been formulated in section 4.2. However, this kind of paradigm occurs far more often than the 'rare' paradigms. Therefore, it is classified as 'semi-common'. The Kwakiutl-type paradigmatic structure has both a horizontal homophony between 3 and 3+3 and one between 2 and 2+3. An example of this type is found in Kwakiutl, a Wakashan language from Canada. Shown in (4.84) are the subject suffixes from Kwakiutl (Boas, 1947:252).⁴²

⁴² The system for spacial deixis is disregarded here. I interpret this spatial system as a separate system, although it cross-cuts through the pronominal system. The same structure as in Kwakiutl is found in the related language Heiltsuk (Rath, 1981:77).

(4.84) KWAKIUTL

		...-ents	<i>I</i> +2
		...-enu ^s _g ^u	<i>I</i> +2+3
<i>I</i>	...-en(L)		<i>I</i> +3
<i>2</i>	...-εs		<i>2</i> +3
<i>3</i>	...-∅		<i>3</i> +3

Three other examples of paradigms of the Kwakiutl-type are shown below, all from South America. First, the intransitive prefixes from Apalai, a Carib language from Brazil, are shown in (4.85). The (non-collective animate) independent pronouns have the same paradigmatic structure (Koehn & Koehn, 1986:95, 108). Next, the intransitive suffixes from Huallaga Quechua, a Quechuan language from Peru, are shown in (4.86); data from Weber (1986:334). Finally, the intransitive subject pronouns from Maxakali, a Gé language from Brazil, are shown in (4.87). Other pronominal paradigms in Maxakali have an identical structure (Popovich, 1986:352-353, 356). It is noted that ‘the particles under discussion are separate words, except where they are phonologically bound to the preceding words by fusion’ (Popovich, 1986:351). It is difficult what to make of this statement; I will consider the Maxakali pronominal morphemes as independent elements.

(4.85) APALAI

		s(y)-...	<i>I</i> +2
		ynan(y)-...	<i>I</i> +2+3
<i>I</i>	∅/y-...		<i>I</i> +3
<i>2</i>	o/m-...		<i>2</i> +3
<i>3</i>	n(y)-...		<i>3</i> +3

(4.86) HUALLAGA QUECHUA

		...-nchi:	<i>I</i> +2
		...-:kuna	<i>I</i> +2+3
<i>I</i>	...-:		<i>I</i> +3
<i>2</i>	...-nki		<i>2</i> +3
<i>3</i>	...-n		<i>3</i> +3

(4.87) MAXAKALI

		yũmũg	<i>I</i> +2
		‘ũgmũg	<i>I</i> +2+3
<i>I</i>	‘ũg		<i>I</i> +3
<i>2</i>	‘ã		<i>2</i> +3
<i>3</i>	‘ũ		<i>3</i> +3

I know of two other examples of a Kwakiutl-type pronominal paradigm outside America. First, the ‘object’ prefixes from Svan, a Caucasian language from Georgia, are shown in (4.88). The separate exclusive pronoun is probably an innovation in Proto-Svan (Tuite, 1997:23). Second, the independent pronouns from Acehnese are

presented in (4.89). Note that there are different honorific versions of these pronouns. The presented morphemes are the ‘familiar’ forms (Durie, 1985: 117).

(4.88) SVAN

		gw-...	<i>1+2</i>
		n-...	<i>1+2+3</i>
<i>1</i>	m-...		<i>1+3</i>
<i>2</i>	ž-...		<i>2+3</i>
<i>3</i>	x-...		<i>3+3</i>

(4.89) ACEHNESE

		geutanyoe	<i>1+2</i>
		kamoe	<i>1+2+3</i>
<i>1</i>	kee		<i>1+3</i>
<i>2</i>	kah		<i>2+3</i>
<i>3</i>	jih		<i>3+3</i>

4.6.5 Rare variants of ‘inclusive/exclusive’

Only a few instances of other paradigmatic structures with specialised inclusive and exclusive marking are attested. All these ‘rare’ examples have some kind of vertical homophony. Three different kinds of vertical homophony are found: between the inclusive and 2+3, between the exclusive and 3+3 and between 2+3 and 3+3. The most striking aspect is that there are only so few cases. Surely, when compared with the large amount of vertical homophonous cases that were found among the ‘unified-we’ paradigms (41 cases), the seven examples attested make a real exceptional impression.

Tiwi-type paradigm

The first set of examples all have a vertical homophony between the inclusive and the second person plural. The first case is the object marking in the Australian language Tiwi, shown in (4.90); data from Osborne (1974: 39).

(4.90) TIWI

		mani-...	<i>1+2</i>
		məwəni-...	<i>1+2+3</i>
<i>1</i>	məni-...		<i>1+3</i>
<i>2</i>	məni-...	mani-...	<i>2+3</i>
<i>3</i>	∅-...	wəni-...	<i>3+3</i>

Sanuma-type paradigm

The second case is found in Sanuma, a Yanomam language from Venezuela/Brazil. Shown in (4.91) is the paradigm of the non-emphatic independent pronouns (Borgman, 1990: 149). This paradigm is almost like the one from Tiwi, except for the horizontal homophony in the third person. The unusual vertical homophony between ‘we’ and ‘you’ is exemplified in sentence (4.92a). There is a special exclusive ‘we’ morpheme. This exclusive form is exemplified in sentence (4.92b). In this sentence, the speaker is telling about a group of non-present co-workers, using an exclusive ‘we’ indicating that the addressee listening to the story was not part of that group.

(4.91) SANUMA

			makö	1+2
			samakö	1+2+3
1	sa		samakö	1+3
2	wa		makö	2+3
3	(classifiers)			3+3

(4.92) SANUMA

- a. **makö** *kali-palo* *mai* *kite*
 1+2/2+3 work-REPET NEG FUT
 ‘We/you (Plur) are not going to work’ (Borgman, 1990: 150)

- b. *Poa Pisita* *ha* *ĩ,* **samakö** *pewö* *kali-palo*
 PLACE LOC REL 1+3 all work-REPET
 ‘In Boa Vista we all work’ (Borgman, 1990: 153)

Nama-type paradigm

The third case of a vertical homophony between the inclusive and the second person plural is found in Nama. The pronouns of the Khoisan language Nama consist of two parts, which Hagman (1977:43) calls the ‘pronominal root’ and the ‘png-suffix’. Haacke (1977) argues even stronger that these two parts have a completely different status in the grammar of the language. Following Haacke, I consider them to be two different pronominal paradigms. The ‘pronominal root’ has a vertical and a horizontal homophony, shown in (4.93). The inclusive reference is identical to the second person singular and plural (Hagman, 1977:44).

(4.93) NAMA

			saá	1+2
			sií	1+2+3
1	tií		sií	1+3
2	saá			2+3
3	//ĩ			3+3

Shuswap-type paradigm

The next examples all have a vertical homophony involving the exclusive (1+3). In these examples, the exclusive is identical to the third person plural. The first example of this homophony is the paradigm of the intransitive suffixes of Shuswap, a Salish language from Canada. In Shuswap, the exclusive is marked identical to the third person singular and non-singular, shown in (4.94). An independent pronoun *k°əx°* is used to disambiguate the exclusive ‘we’ from the third person (Kuipers, 1974:45, 59).

(4.94) SHUSWAP

			...-ət	1+2
			...-əs	1+2+3
1	...-wn		...-əs	1+3
2	...-əx°		...-əp	2+3
3	...-əs			3+3

Waiwai-type paradigm

An almost identical paradigm is attested in the intransitive pronominal prefixes from Waiwai, a Carib language from Brazil, shown in (4.95). Just as in Shuswap, the exclusive is marked identical to the third person singular and non-singular. Again just as in Shuswap, this unusual pattern of Waiwai needs disambiguation. The difference between the category 1+3 and the third person singular and non-singular. The difference between the category 1+3 and the third person singular and non-singular has to be clarified by the use of an independent pronoun *amna* for the exclusive ‘we’. A pair of examples demonstrating the Waiwai equivalents of the exclusive and inclusive are shown in (4.96). Different from Shuswap, the Waiwai pronominal prefixes also show a horizontal homophony in the second person (Hawkins, 1998: 178-179).⁴³

(4.95) WAIWAI

		t(i)-...	1+2
			1+2+3
1	k(i)-...	n(i)/∅-...	1+3
2	m(i)-...		2+3
3	n(i)/∅-...		3+3

(4.96) WAIWAI

- a. *amna nu-puru*
 1+3,PRON 3-roast
 ‘We (exclusive) roasted it’ (Hawkins, 1998: 179)
- b. *tî-hînoyasî amê*
 1+2-know later
 ‘We (inclusive) will know it later’ (Hawkins, 1998: 179)

Kunama-type paradigm

Finally, a third form of vertical homophony is found in Kunama, a Nilo-Saharan language from Eritrea, shown in (4.97). In this paradigm, there is a vertical homophony between the second and third person non-singular. The presented suffixes are used with verbs of type I/type S (Reinisch, 1881: 53; Bender, 1996: 19).

(4.97) KUNAMA

		...-di	1+2
			1+2+3
1	...-na	...-ma	1+3
2	...-nu		2+3
3	...-su	...-mu	3+3

4.6.6 Summary

Two paradigmatic structures with an ‘inclusive/exclusive’ first person complex are attested commonly among the world’s languages (the Mandara-type and the Tupí-Guaraní-type). A third pattern (the Kwakiutl-type) is also attested regularly, but it

⁴³ The same paradigmatic structure as in Waiwai is also found in another Carib languages, Hixkaryana (Derbyshire, 1979: 146-149). In this language, the independent pronoun ‘amna’ is also used to clarify the referent as exclusive.

does not satisfy all criteria that were raised for a paradigm to be classified as ‘common’. This paradigmatic type is ‘semi-common’. The paradigmatic structure of these (semi-)common patterns is outlined in Figure 4.17. These three common paradigmatic structures are reminiscent of the three common structures of ‘unified-we’ (see Figure 4.9 on page 126 above). Those three common ‘unified-we’ patterns are identical to the three patterns as shown here, except for the inclusive-exclusive distinction.

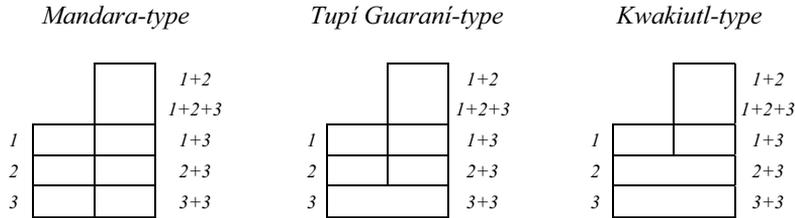


Figure 4.17: (Semi-)common paradigmatic structures of ‘inclusive/exclusive’

Structurally, the Mandara-type paradigms can be interpreted as the ‘complete’ patterns, relative to which the Tupi Guarani-type does not distinguish between singular and non-singular in the third person. Besides, the Kwakiutl-type also neutralises the singular/non-singular distinction in the second person. This hierarchical interpretation of the three structures is substantiated by the relative frequencies of the paradigms, as shown in Table 4.15. The Mandara-type is the most common. The number of examples of the Tupi Guarani-type is already down to roughly half of that. The Kwakiutl-type is even less frequent. These counts are specified in Appendix B.

	<i>number of cases</i>
<i>Mandara-type</i>	38
<i>Tupi Guarani-type</i>	14
<i>Kwakiutl-type</i>	8
Total	60

Table 4.15: Number of cases of the (semi-)common paradigms

Relative to the three (semi-)common paradigmatic structures, only very few other paradigmatic structures are attested. Only seven cases of ‘rare’ paradigmatic structures are attested, relative to 60 cases of the three (semi-)common structures. The distinction between common and rare is rather clear-cut in the case of ‘inclusive/exclusive’ first person complex.

	<i>common</i>	<i>semi-common</i>	<i>rare</i>
<i>examples</i>	52	8	7
<i>paradigmatic structures</i>	2	1	6
<i>examples per structure</i>	26.0	8.0	1.2

Table 4.16: Frequencies of the ‘inclusive/exclusive’ paradigms

4.7 Variants of the ‘minimal/augmented’ type

4.7.1 Preamble

The final kind of paradigmatic structures to be discussed is characterised by a ‘minimal/augmented’ first person complex. This paradigmatic structure has three different forms that are to be translated with the English pronoun ‘we’. These three different forms of ‘we’ are traditionally referred to as ‘dual’, ‘inclusive’ and ‘exclusive’. This traditional tripartite division is renamed here as ‘minimal inclusive’ (1+2), ‘augmented inclusive’ (1+2+3) and ‘exclusive’ (1+3) respectively (see section 3.6.5 for an extensive discussion).

$1+2$			A	A	A
$1+2+3$	A	–			B
$1+3$			–	B	C

Table 4.17: The ‘minimal/augmented’ pattern

Theoretically, many different paradigmatic structures could be made by combining the marking of the ‘minimal/augmented’ first person complex with the remaining categories in the paradigm. However, *pace* two cases, all examples of the ‘minimal/augmented’ type belong to one and the same paradigmatic structure, as presented schematically in Figure 4.18. This kind of paradigm will be called the MARANAO-TYPE paradigm.

			$1+2$
			$1+2+3$
1			$1+3$
2			$2+3$
3			$3+3$

Figure 4.18: The Maranao-type paradigmatic structure

4.7.2 Maranao-type paradigm

The independent pronouns from Maranao, an Austronesian language from the Philippines, are presented in (4.98). The paradigm consists of eight different forms. In this paradigmatic structure, the category 1+2 is referentially a dual form. It refers to the two participants of the speech-act dyad: the speaker and the addressee. However, calling it a ‘dual’ does not seem to be the best analysis. The most important argument against the term ‘dual’ is that paradigms like the one from Maranao do not have any other dual forms. To invoke a completely new category ‘dual’ for the analysis of such a paradigm is a waste of apparatus. The ‘singular-group’ analysis proposed here takes care of the category without any new terminological dimension. All pronominal paradigms that have three different forms for ‘we’, and that do not have any second or third person dual or trial forms, are classified as minimal/augmented paradigms. Notwithstanding the plethora of descriptive terminology that is found in grammars to describe this kind of paradigm, I have decided to take them all together and consider

them as one and the same type. It then turns out that this seemingly exotic paradigmatic structure is found relatively frequent among the world's languages.

(4.98) MARANAO

		ta	1+2
		tano	1+2+3
1	ako	kami	1+3
2	ka	kano	2+3
3	sekanian	siran	3+3

The description of Maranao-type paradigms without the use of the term 'dual' started with the analysis of some Philippine languages in the fifties (Thomas, 1955). It is not a surprise that this analysis first arose with the description of a language in this part of the world, as paradigms of the Maranao-type are abundantly found on the Philippines. Well known cases are Maranao (McKaughan, 1959) and Hanunóo (Conklin, 1962:134-135). A large number of examples of this paradigmatic structure are described in a collection edited by Reid (1971:1-43). This author presents short descriptions of the phonology and pronominal categories of 43 'minor' Philippine languages. Of these 43 languages, 31 are described as having a Maranao-type paradigm. The other 12 languages have only one inclusive form, and are classified as instances of 'inclusive/exclusive' paradigms.⁴⁴ Both types of languages occur distributed equally over the geographical area of the Philippines and through all linguistic (sub)families. Reid (1971:x) hints in his introduction that the languages without the opposition between the two different inclusive categories originally had a paradigm of the Maranao-type, but lost the opposition between 1+2 and 1+2+3. Examples of Maranao-type paradigms are attested frequently among the non-Pama-Nyungan languages from Australia.⁴⁵ It is found in the inflection of the Daly river languages. Properly speaking, it is not the verb itself that is inflected, but the so-called auxiliary verb that always accompanies a verbal stem.⁴⁶ A Maranao-type paradigm is also attested in the independent pronouns of the Daly language Malakmalak. Note that these pronouns make a gender distinction in the third person singular (Birk, 1976:30-31). The same situation, a Maranao-type paradigms with gender in the third person singular, is described for the independent pronouns of Tiwi (Osborne, 1974:54). Some more exam-

⁴⁴ The following names of languages are taken from the collection by Reid (1971:1-43). The names between brackets are the names of a subvariety of a certain language, normally an indication of the location where it is spoken. A difference between 1+2 and 1+2+3 is found in Agta, Atta (Pamplona), Balangaw, Batak (Palawan), Bilaan (Korondai and Sarangani), Binukid, Bontoc (Guinaang), Dumagat (Casiguran), Ifugao (Amganad, Batad and Bayninan), Inibaloi, Isneg, Itneg (Binongan), Kalagan, Kalinga (Guinaang) Kallahan (Kayapa Proper and Keleyqiq), Kankanay (Northern), Manobo (Dibabawon, Ilianen, Sarangani and Western Bukidnon), Samal, Sambal (Botolan), Sangil (Sarangani Islands), Sangir, Tagabili, Tagbanwa (Aborlan) and Tausug. An identical morpheme for 1+2 and 1+2+3 is found in Gaddang, Ilongot (Kakiduge:n), Itbayaten (Batanes Islands), Ivatan (Batanes Islands), Mamanwa, Manobo (Ata, Kalamansic Cotabato and Tigwa), Mansaka, Subanun (Sindangan), Subanon (Sioccon) and Tagbanwa (Kalamian).

⁴⁵ Dixon (1980:351-356) speculates whether a Maranao-type paradigms might be reconstructed for Proto-Australian, but this seems implausible. A more fruitful proposal is to reconstruct a Maranao-type paradigm only for the non-Pama-Nyungan languages (Blake, 1988:7; 1991:222).

⁴⁶ Eg Maranungku (Tryon, 1970:17-42) and Malakmalak (Birk, 1976:47-81). For a wider view on the Daly river auxiliary verbs and their person suffixes, see Tryon (1976:679).

ples of the Maranao-type paradigm in Australia are found among the independent pronouns of the Nyulnyulan languages.⁴⁷ An incidental case of a Maranao-type paradigm is attested in southeastern New Guinea. The independent pronouns from Papuan language Mountain Koiali show this paradigmatic structure. Other related Koiarian languages do not have such a paradigm (Garland & Garland, 1975:429,434-435). A final example in the Pacific of a Maranao-type paradigm is countered in the East Papuan languages Santa Cruz and Nanngu, spoken in Solomon Islands (Wurm, 1969:77-83). Quite unexpectedly, the Maranao-type paradigm shows up rather frequently in Africa. All examples are found in the ‘elbow’ of Africa, mainly in Cameroon and Nigeria. The Maranao-type occurs in both in the Niger-Congo stock and in the Chadic family (Afro-Asiatic stock).⁴⁸ I know of no examples from the Nilo-Saharan stock. Finally, the Maranao-type paradigm is also found in some languages from the West Coast of the USA (California and Oregon).⁴⁹ In this region, some examples of inflectional paradigms of the Maranao-type are found. Still, most examples of the Maranao-type seem to be independent pronouns.

4.7.3 Rare variants of ‘minimal/augmented’

Only two cases are attested that distinguish the three different forms for ‘we’ (as in the Maranao-type paradigm), but that do not have precisely the same paradigmatic structure. These exceptional paradigms are found in Ilocano and in Tiwi.

Ilocano-type paradigm

The first explicit discussion of a Maranao-type pronominal system was by Thomas (1955) for the Philippine language Ilocano. The independent pronouns of Ilocano are shown in (4.99). In his article, Thomas criticised the traditional analysis with a ‘first person dual inclusive’ as name for the category 1+2. He also criticised the categorisation of *na* and *da* as third person pronouns. He concluded that: ‘it is readily seen that *na* and *da* are not true pronouns but may be simply number indicators’ (Thomas, 1955:207).

⁴⁷ Eg Bardi (Metcalfe, 1975:48-50) and Nyulnyul (McGregor, 1996:23).

⁴⁸ Among the Niger-Congo languages, the Maranao-type paradigm is attested in the basic pronouns of the Grassfields languages in Cameroon, for example Limbum (Fransen, 1995:179-180), Bamileke (Anderson, 1985:63,68) and Babungo (Schaub, 1985:193-194). It is also found in Dii, an Adamawa language from Cameroon (Bohnhoff, 1986:126-127) and in Dan, a Mande languages from Côte d’Ivoire (Doneux, 1968:45-47). Among the Chadic languages, the Maranao-type paradigm is found in the independent pronouns from Marghi (Hoffmann, 1963:73-74; Burquest, 1986:77,82) and Gude (Hoskison, 1983:48). The Margi pronouns seem to cliticise when occurring behind the verb.

⁴⁹ A Maranao-type paradigm is attested in the independent pronouns and in the pronominal prefixes of the northern languages of the Uto-Aztecan family. It is described extensively for Ute-Southern Paiute. However, there are gender oppositions and special spatial forms (visible versus invisible) in the third person (Sapir, 1930:176-178; Press, 1975:849-90; Givón, 1980:50; 1984:356). The same paradigmatic structure is also found in Tübatulabal, another language of the Northern Uto-Aztecan branch. In the description of Tübatulabal by Voegelin (1937a:135), the exclusive form *-(g)ila’ay* is glossed as ‘dual exclusive’, but this is a mistake. This morpheme can also have plural reference. This can be inferred from the usage of this morpheme in text 27, especially in a scene where a group of 5 or 6 boys are playing (Voegelin, 1937b:225, sentences 116-134). The narrator, who was one of the boys, tells the audience that ‘we were playing during the day’ using the exclusive form. This form clearly does not have dual reference in this context. Outside the Uto-Aztecan family, but areally in close contact, a Maranao-type paradigm is found in the independent pronouns and in the intransitive pronominal prefixes of Southern Sierra Miwok (Broadbent, 1964:43,93).

(4.99) ILOCANO

		ta	$1+2$
		tayo	$1+2+3$
1	co	mi	$1+3$
2	mo	yo	$2+3$
3	(number markers)		$3+3$

Tiwi-type paradigm

The only other variant on the minimal/augmented theme is attested in the Australian language Tiwi. The intransitive subject prefixes in the non-past are presented in (4.100). The precise function of the affix ‘pə’ is unclear. It is, among other things, conditioned by tense and signals gender of the referent in the third person singular (Osborne, 1974:38).

(4.100) TIWI

		mu-...	$1+2$
		ŋa-...	$1+2+3$
1	ŋə-...	ŋə-pə-...	$1+3$
2	ŋə-pə-...	wu-...	$2+3$
3	a(pə)-...		$3+3$

4.7.4 Summary

The examples of the minimal/augmented type show almost no variation in paradigmatic structure. By far the majority of examples are of the Maranao-type paradigm. In this paradigmatic structure, all eight categories are distinguished by different morphemes.

			$1+2$
			$1+2+3$
1			$1+3$
2			$2+3$
3			$3+3$

Figure 4.19: The Maranao-type paradigm

Only two other paradigmatic structures are attested in incidental cases. The frequencies of the cases as discussed in this section are presented in Table 4.18. These numbers are accounted for in Appendix B. The distinction between common and rare is rather clear-cut in the case of ‘minimal/augmented’ paradigms.

	<i>common</i>	<i>rare</i>
<i>examples</i>	22	2
<i>paradigmatic structures</i>	1	2
<i>examples per structure</i>	22.0	1.0

Table 4.18: Attested frequencies of the ‘minimal/augmented’ types

4.8 Conclusion

In this chapter, a wide variety of possible paradigmatic structures has been presented. In total, 55 different structures have been described. Together with the 6 structures from the previous chapter, a total of 61 pronominal paradigms has been found (all these counts are accounted for in Appendix B). From this impressive variation, it can already be inferred that the traditional six-way paradigm (ie first, second and third person in the singular and plural) is not the only and probably also not the basic paradigmatic structure for a pronominal paradigm. It seems rather far fetched to proclaim all other kinds of paradigms to be either extended or corrupted versions of the six-way paradigm. Accordingly, one of the major goals of this chapter has been to establish which other paradigmatic structures are commonly found among the world's languages and which ones are the 'real' exceptions amidst the massive variation. A sample has been constructed on the basis of the description of the diversity from this chapter. This sample has been composed to perform all kinds of quantitative analyses. The method to construct the sample and the results of the analyses will be described extensively in the next chapter. However, this sample is used here to give an indication of the frequency of the various paradigmatic structures. The distribution of the total sample of 265 cases over the 61 different paradigms is presented in Table 4.19.

	<i>common</i>			<i>semi-common</i>			<i>rare</i>		
	<i>C</i>	<i>P</i>	<i>C/P</i>	<i>C</i>	<i>P</i>	<i>C/P</i>	<i>C</i>	<i>P</i>	<i>C/P</i>
<i>unified-we</i>	67	3	22.3	22	6	3.7	19	15	1.3
<i>no-we</i>	20	1	20.0				16	11	1.5
<i>only-inclusive</i>	15	1	15.0				7	6	1.2
<i>inclusive/exclusive</i>	52	2	26.0	8	1	8.0	7	6	1.2
<i>minimal/augmented</i>	22	1	22.0				2	2	1.0
<i>others</i>							8	6	1.3
Total	176	8	22.0	30	7	4.3	59	46	1.3

Table 4.19: Summary of frequencies, as counted in Appendix B

(C = number of cases, P = number of paradigmatic structures,
C/P = cases per paradigmatic structure)

Nine different paradigms were characterised as 'common', with a mean of 22.0 cases per paradigm. The remaining cases are clearly less common, although seven paradigmatic structures still occur relatively frequent. These seven paradigms were characterised as 'semi-common', with a mean occurrence of 4.3 cases per paradigm. The remaining 46 paradigms are clearly rare, with a mean occurrence of 1.3 cases per paradigm.

The distinction between the 'common' and 'rare' has been made on partly qualitative grounds. The common paradigmatic structures are attested often, widely dispersed over the world's languages and they are a general characteristic of at least a few genetically close-knit groups. In contrast, the rare paradigmatic structures are only found in incidental cases; even within their close family the same paradigmatic structure is

normally not found. In the sample, the distinction between ‘common’ and ‘rare’ is a continuum, mediated by the ‘semi-common’ paradigmatic structures. The continuum is presented in Figure 4.20, showing the various paradigmatic structures from most frequent to less frequent (only three of the 36 paradigmatic structures that are represented by one example are shown; the periods on the top of the figure indicate that there are many more cases). From this continuum, it might be inferred that the division in ‘common’, ‘semi-common’ and ‘rare’ is rather arbitrarily chosen. However, I will argue that this division presents some interesting generalisations over the structure of the paradigms.

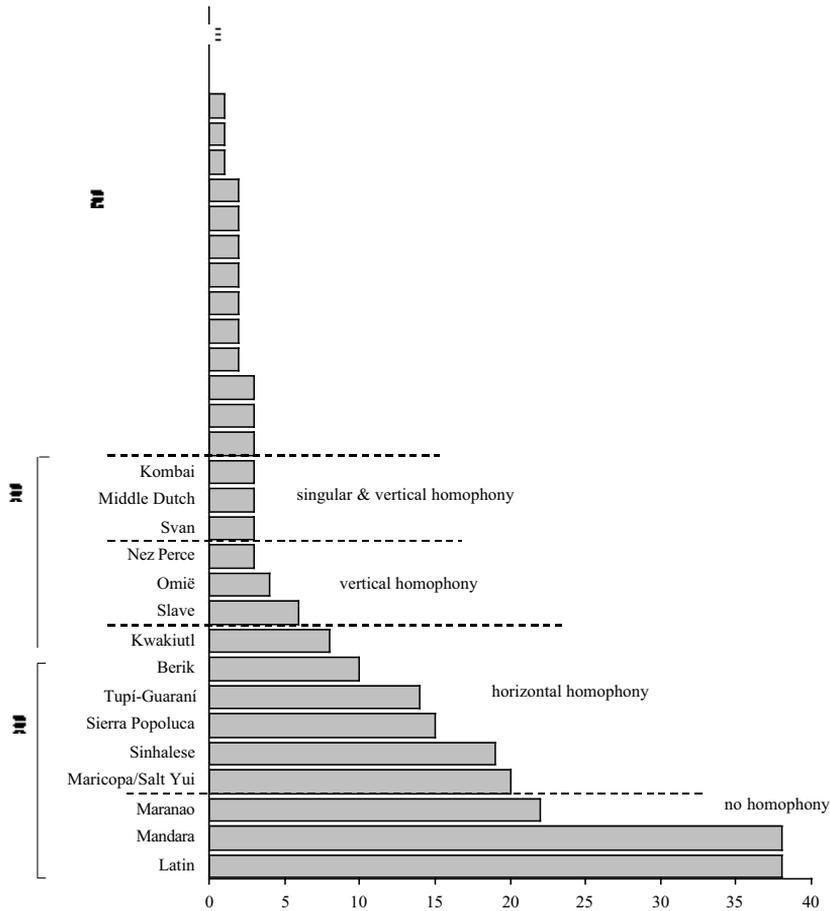


Figure 4.20: Continuum of occurrence of paradigmatic structures
(Only three of the 36 paradigmatic structures that are
represented by one case are shown in the figure)

The Latin-type and the Mandara-type paradigm, the two most frequent paradigmatic structures, were characterised by Ingram (1978:219) as ‘more frequent than the others’. Indeed, these two are most frequent, but they are not qualitatively different from the others. They are better seen as part of a continuum leading from more to less fre-

quent. When the actual structures on this continuum are scrutinised, four qualitative breaks can be discerned. The first qualitative break occurs after the first three paradigmatic structures: these three structures are all ‘full’, ie they do not show any horizontal, vertical or singular homophony. These three paradigms are shown in the left-most column in Figure 4.21.

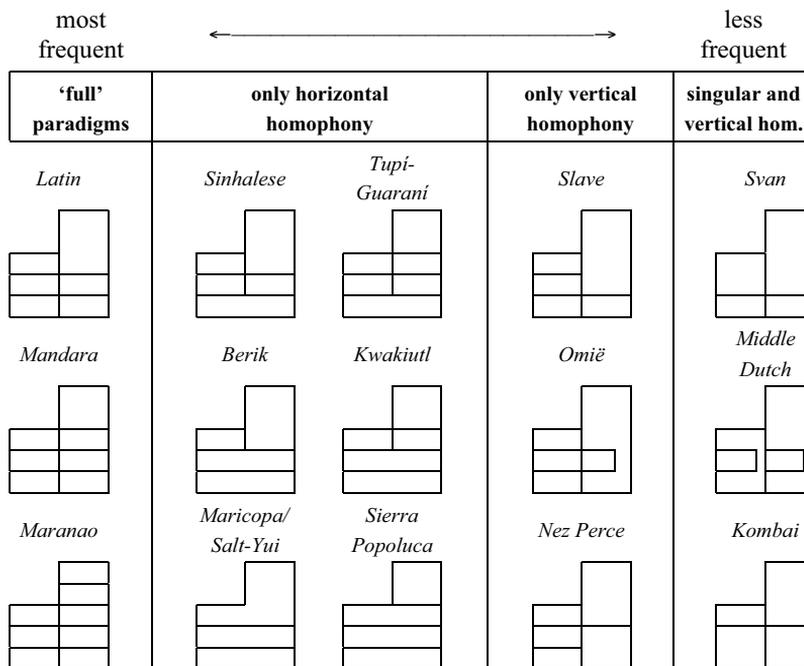


Figure 4.21: Frequency continuum of the common and semi-common paradigmatic structures, ordered to structural characteristics

The second qualitative break occurs after the next six paradigms, roughly between the ‘common’ and ‘semi-common’ structures. These six paradigmatic structures have only horizontal homophony. The kind of horizontal homophony follows a hierarchy $3 > 2 > 1$. These six paradigms are shown in the second column in Figure 4.21.⁵⁰ The third qualitative break is found after the next three paradigms, amidst the ‘semi-common’ paradigms. The three most frequent of these ‘semi-common’ paradigms only show vertical homophony (shown in the third column of Figure 4.21). The final qualitative break is found again after three paradigms, dividing ‘semi-common’ from ‘rare’. The three remaining ‘semi-common’ paradigms have a vertical and a singular homophony, which are mirror-images of each other. The remaining ‘rare’ paradigmatic structures show all kind of combinations of the various kind of homophony.⁵¹

⁵⁰ The Kwakiutl-type paradigm is ‘semi-common’, but the structure of this paradigm belongs to the other paradigms before the second qualitative break, together with the other ‘common’ types of paradigms.

⁵¹ The three paradigmatic types that follow on the frequency-continuum are the English-type, the Ainu-type and the Aztec type paradigm. These paradigms have only horizontal homophony, and no vertical

This quick qualitative analysis presents a good prospect for the detailed structural analyses that will follow in the next chapter. There seems to be a hierarchy in the kind of horizontal homophony, third person above second above first. Also, the various kinds of homophony seem to be kept rather distinct: horizontal homophony occurs most likely without vertical or singular homophony. In contrast, singular homophony is most likely to occur together with vertical homophony. The fine details of these trends will be discussed in the next chapter.

[continued from previous page]

or singular homophony. The English-type has a homophony between 2 and 2+3 (see page 114). The Ainu-type has precisely the opposite structure of the English-type paradigm: the categories 2 and 2+3 do not show horizontal homophony (see page 134). Finally the Aztec-type has a ‘diagonal’ homophony between the first person complex and the second person singular (see page 134). I do not see any interesting generalisation in this set of paradigmatic structures.

Chapter 5

Paradigmatic explicitness

Structural analysis of oppositions in the person paradigms

5.1 Introduction

In this chapter, the perspective changes from a ‘paradigmatic view’ into a ‘categorical view’. In the previous chapter, the pronominal paradigms were presented in their entirety; a perspective that I will return to in subsequent chapters. However, in this chapter, the INDIVIDUAL MORPHEMES in the paradigm and their referential properties will be scrutinised. A pronominal paradigm is interpreted as a sort of micro-‘Wortfeld’ (Trier, 1931) in which the meaning of the individual items can only be characterised in opposition to the other items in the paradigm. Various person-oppositions that are found in pronominal paradigms turn out to be interrelated in the sense that, cross-linguistically, the occurrence of specific oppositions is strongly predicted by the occurrence of other oppositions in the paradigm. It will be shown that there are two different complexes of such interrelated oppositions, ordered along the lines of two different hierarchies. Both these hierarchies are gradients of EXPLICITNESS of a pronominal paradigm. This amount of explicitness of a pronominal paradigm in turn correlates with the morphological status of the paradigm. There is a strong tendency for the more explicit paradigms to be marked as independent pronouns and for the less explicit paradigms to be inflectionally marked.

5.2 Method and terminology

The data for the quantitative analyses to be presented are formed by the collection of pronominal paradigms as described in the previous chapter (a summary of the crucial characteristics of all cases can be found in Appendix B). In that chapter, 61 different paradigmatic structures were presented. These cases included pronominal paradigms that only grammaticalise singular and general non-singular (‘group’) marking. The discussion of paradigms with specialised number marking (like dual and paucal) is postponed to Part Three. Paradigms with sub-categorisations (like gender or honorifics) were ignored, except for a few incidental cases with gender marking in the third person singular. A sample has been selected of 265 paradigms on the basis of the paradigmatic variation as presented in the previous chapter. The collection of the 265 cases has been guided by the wish to include the complete variation known and thus to show the inherent variability of linguistic structure. In principle, all examples that were described in the previous chapter are included in the sample. This means that all

61 different paradigmatic structures, including all quaint cases, are represented in the sample. The only constraint that has been added on the construction of the sample is that genetic families never count for more than four cases. This restriction has been introduced to avoid overrepresentation of paradigmatic structures that are commonly found in large and widespread families. Because of this approach, the present ‘diversity sample’ is not completely balanced over the world’s linguistic diversity as measured by quantity and genetic distance (cf Rijkhoff & Bakker, 1998). However, it is not at all clear *a priori* whether genetic diversity is an important factor for paradigmatic diversity. For example, in the next chapter it will be argued that there is much paradigmatic variation within close genetic distances. Of course, a cross-linguistic sample should draw data from the known linguistic diversity of the world’s languages. In other words, genetic diversity should be taken as a basis for the construction of a cross-linguistic sample. However, a tight method of genetic sampling only reduces the possibility to highlight interesting cases. In the present collection, the number of cases and the distribution of these cases of the world’s languages is wide enough to grant a cross-linguistic interpretation. The 265 cases are taken as a sample of the world’s languages on the basis of which I will investigate what kind of oppositions are found in which quantity in pronominal paradigms.

In practice, this investigation is carried out by making an inventory of the attested combinations of the eight person categories, as developed in Part One. A maximally explicit paradigm has oppositions between all eight referential categories, as shown in Figure 5.1. Most paradigms do not mark all different categories by different morphemes. These paradigms combine various categories into the referential properties of one morpheme. In this chapter, I will show which combinations are found more often than others and what kinds of interdependencies exist between the various combinations.

		<i>‘non-singular’</i>			
<i>‘singular’</i>		1+2	<i>minimal inclusive</i>	}	<i>inclusive</i>
		1+2+3	<i>augmented inclusive</i>		
<i>speaker</i>	1	1+3	<i>exclusive</i>	}	
<i>addressee</i>	2	2+3	<i>second person plural</i>		
<i>other</i>	3	3+3	<i>third person plural</i>		
				<i>first person complex</i>	

Figure 5.1: Maximally explicit person marking

In this chapter, I will focus on three different kinds of combinations of the eight pronominal categories, which will be called SINGULAR, VERTICAL and HORIZONTAL homophony. The three kinds of homophony are illustrated in Figure 5.2. This figure only presents illustrative examples of the three kinds of homophony. There are many more combinatorial possibilities for each kind of homophony. The term ‘homophony’ is only intended as a descriptive device to refer to certain kinds of combinations of referential categories into one morpheme. I do not have any theoretical assumptions as to which kind of homophony is ‘normal’ and which kind is ‘abnormal’. The quantitative analyses will show whether such a division between normal and abnormal is justified, and if so, which kinds of homophony are more ‘normal’ than others. The

(5.2) Person Hierarchy (two-stage version):

speaker, addressee > other

speaker > addressee

The first stage of this hierarchy is used to derive a person hierarchy for non-singular reference, as shown below in (5.3). Categories that include reference to both speaker and addressee rank above categories that include only reference to either speaker or addressee. This, in turn, outranks categories that include reference to neither speaker nor addressee.

(5.3) Person Hierarchy (non-singular version):

speaker AND addressee > speaker OR addressee > NEITHER speaker NOR addressee

Finally, this hierarchy is combined with the ranking of speaker above addressee to form a complete person hierarchy, as shown in (5.4). This hierarchy turns out to govern the occurrence of various forms of homophony, as will be shown in this chapter. The same hierarchy is also found to govern ‘dualworthiness’ (Plank & Schellinger, 2000).

(5.4) Person Hierarchy (complete version):

speaker AND addressee > speaker > addressee > NEITHER speaker NOR addressee

Equipped with this terminological and conceptual framework, the journey through the data can begin. First, I will turn to the different forms of vertical homophony. There is a large variety of this kind of homophony, but not all variants are equally ubiquitous (section 5.3). Next, singular homophony will be discussed. A complete description of the various forms of singular homophony has already been presented in chapter 2. Here, I will only show that the occurrence of singular homophony is highly constrained by various other characteristics of the pronominal paradigm (section 5.4). Then, the first large generalisation over the structural oppositions in pronominal paradigms is presented: the Explicitness Hierarchy. This hierarchy brings together vertical and singular homophony, together with the various marking patterns of the first person complex (section 5.5). The occurrence of horizontal homophony shows a different pattern, although its occurrence is also highly constrained. These constraints make up the Horizontal Homophony Hierarchy, which turns out to have interesting connections with the Explicitness Hierarchy (section 5.6). An important typological cut-off point in these hierarchies is the presence of absence of an inclusive/exclusive opposition. The effect of this parameter will be presented in the form of a semantic map (section 5.7) Both hierarchies are correlated with the morphological status of the pronominal elements: the more explicit the marking, the larger the proportion of independently marked paradigms (section 5.8). Finally, the main findings of this chapter are brought together in a concluding section (section 5.9).

5.3 Vertical Homophony**5.3.1 Preamble**

Vertical homophony is characterised by the fact that the categories 2+3 (‘second person plural’) or 3+3 (‘third person plural’) are marked identically to any of the first person complex categories, or identically to each other. Theoretically, many different

kinds of vertical homophony are possible. If there is no vertical homophony (ie there are two separate morphemes that mark for 2+3 and 3+3 and both morphemes do not mark for any other non-singular category), then the paradigm is called ‘split’. A SPLIT NON-SINGULAR paradigm is the opposite of a VERTICAL HOMOPHONOUS paradigm.

To exemplify this definition, two rather arbitrarily chosen examples of vertical homophony are presented below. First, the independent pronouns of the Slavey dialect of Slave, an Athabascan language spoken in Canada, are presented in (5.5). In this paradigm, the category 2+3 is identical to the unified first person complex, resulting in a homophonous form *nax̣i* that is used with the equivalent reference to the combination of the English *we* and *you-plural* (Rice, 1989:253,431).

(5.5) SLAVE

1	ṣị	nax̣ị	1+2
2	ṇị		1+2+3
3	ʔeḍị		1+3
		ʔegeḍị	2+3
			3+3

The second example of a vertical homophony is attested in the pronominal suffixes from Shuswap, a Salish language from Canada. In this paradigm, the category 3+3 is identical to the exclusive first person plural. The category 3+3 is also identical to the singular category 3. This is a case of horizontal homophony, a subject that will be discussed in section 5.6 (Kuipers, 1974:45,59).

(5.6) SHUSWAP

		...-ət	1+2
		...-əs	1+2+3
1	...-wn	...-əs	1+3
2	...-əx̣°	...-əp	2+3
3	...-əs		3+3

Among the 265 cases that form the sample for the present discussion, there are 55 cases with vertical homophony. These 55 cases are found among various paradigmatic structures. However, amidst the variety, the 55 cases have some interesting correspondences. None of these correspondences are universal characteristics; all generalisations that will follow in this section (and in subsequent sections) have exceptions. Still, there are some strong statistical effects that can not be left unnoticed.

5.3.2 Some restrictions on vertical homophony

The first notable effect that has been found is between vertical homophony and the inclusive/exclusive opposition. As can be seen in Table 5.1, there is a strong inverse correlation between vertical homophony and the inclusive/exclusive opposition: if there is a vertical homophony in the paradigm, then there is a strong preference not to mark an inclusive/exclusive opposition. The distribution as presented in this table can be summarised by an implication ‘vertical homophony → no inclusive/exclusive opposition’.

		<i>Inclusive/Exclusive</i>		
		-	+	
<i>Vertical Homophony</i>	-	103	107	210
	+	41	14	55
		144	121	265

Table 5.1: Implication ‘vertical homophony \rightarrow no inclusive/exclusive opposition’
(Fisher’s exact $p = .001$, $\varphi = -.208$)

Another notable correlation holds between vertical homophony and the morphological status of the paradigm. As can be seen in Table 5.2, there is a strong correlation between the presence of a vertical homophony in the paradigm and the paradigm being inflectionally marked. The distribution can be summarised by an implication ‘vertical homophony \rightarrow inflectionally marked’.

		<i>Inflectional</i>		
		-	+	
<i>Vertical Homophony</i>	-	105	105	210
	+	12	43	55
		117	148	265

Table 5.2: Implication ‘vertical homophony \rightarrow inflectionally marked’
(Fisher’s exact $p = .000$, $\varphi = .230$)

The numbers that were shown in the cells of the two previous tables are almost identical. In both tables, the cases without a vertical homophony are evenly distributed; the cases with a vertical homophony are divided roughly in a ration of one to three. However, this correspondence between the two tables is purely due to chance. The two intersecting dimensions, ‘inclusive/exclusive’ and ‘inflectional’, are differently distributed over the languages with a vertical homophony as opposed to the ones without. I will discuss the relation between inflection and the inclusive/exclusive opposition first for the cases with a vertical homophony; followed by the ones without a vertical homophony.

This distribution of the 55 cases WITH A VERTICAL HOMOPHONY over the two dimension ‘inflectional’ and ‘inclusive/exclusive’, as presented in Table 5.3 below, looks like that Holy Grail of the typological textbooks: a tetrachoric table with one (almost) empty cell (Croft, 1990:47-49). This distribution of the 55 cases might suggest an implication ‘inclusive/exclusive opposition \rightarrow inflectional’. However, this is an unwarranted inference. One of the cells in the table is indeed almost empty, but the expected value for this cell is also very low. Only 25.4% of the 55 cases have an inclusive/exclusive opposition, and only 21.8% of the 55 cases is not inflectional. This results in a chance of 5.5% ($0.254 \cdot 0.218 = 0.055$) for the combination inclusive/exclusive and no inflection. Computing 5.5% of the total of 55 cases results in an expected frequency of three cases for this combination. The actually attested two

cases (in the lower left cell of the table) are thus not a sign of an almost empty cell. Two cases are roughly just as much as would be expected from the already intrinsically skewed distribution of the two parameters. The intersection of the two parameters does not show any supplementary interaction between an inclusive/exclusive opposition and inflectional marking (as can also be inferred from the far from significant Fisher's exact: $p = .709$).

		<i>Inflectional</i>		
		–	+	
<i>Inclusive/ Exclusive</i>	–	10	31	41
	+	2	12	14
		12	43	55
		21.8%	78.2%	

Table 5.3: For the 55 cases *with* a vertical homophony, no correlation is attested between inflectional marking and an inclusive/exclusive opposition (Fisher's exact $p = .709$, $\varphi = .107$)

In contrast, for the 210 cases WITHOUT A VERTICAL HOMOPHONY, a (slight) correlation is attested between inflectional marking and the absence of an inclusive/exclusive opposition in the paradigm. At first glance, the distribution in Table 5.4 is not the kind of distribution a typologist would find interesting, as all possible combinations of features are attested in rather large numbers. Still, this distribution shows a significant correlation between inflectional marking and the absence of an inclusive/exclusive opposition. This correlation can intuitively be understood by noting the consistently higher frequencies in the one diagonal as opposed to the other diagonal.¹

		<i>Inflectional</i>		
		–	+	
<i>Inclusive/ Exclusive</i>	–	42	61	103
	+	63	44	107
		105	105	210

Table 5.4: For all 210 cases *without* a vertical homophony, a (slight) inverse correlation between inflectional marking and an inclusive/exclusive opposition is attested (Fisher's exact $p = .013$, $\varphi = -.181$)

This (slight) inverse correlation between inflectional marking and an inclusive/exclusive opposition is not strong enough to warrant any meaningful theoretical

¹ In the typological literature, a (near) zero cell in a tetrachoric table has been automatically interpreted as an implicational universal (Croft, 1990:47-49). However, contrary to this widespread kind of typological reasoning, nothing can be inferred from the distribution as shown in Table 5.3, even though one of the cells is almost empty. In contrast, the distribution from Table 5.4 presents something noteworthy, although all cells are filled with many examples. The observation of an (almost) empty cell in a tetrachoric table is not enough to warrant an implication: the distribution of the other three cells has also to be taken into account. This case shows that an empty cell in a tetrachoric table alone does not suffice to argue for an implicational universal: a significant correlation is also necessary.

consideration. It only hints at a deeper level of analysis that might lead to interesting insights (see section 5.8 for an extensive discussion). For now, the main conclusion to bear in mind is that the 55 cases *with* a vertical homophony behave rather differently from the 210 cases *without* a vertical homophony. This difference can be captured by two implications, repeated here in (5.7) and (5.8).

(5.7) vertical homophony → no inclusive/exclusive opposition

(5.8) vertical homophony → inflectional marking

These implications are to be read as summaries of the attested distributions, not as universal characteristics of human language. The implications summarise the observation that vertical homophony is restricted by two parameters: inflectional marking and the absence of an inclusive/exclusive opposition. Both these restrictions will be used in subsequent sections to develop a theoretical framework to explain the structure of pronominal paradigms. Anticipating the complete framework, the implications can roughly be explained by noting that vertical homophony is a relatively rare phenomenon that will mainly occur on condition that the status of the addressee is not consistently marked through the paradigm (*viz* there is no inclusive/exclusive opposition) and on condition that the salience of the paradigm for the language user is relatively low (*viz* it is inflectionally marked). I will return at length to these explanations in the following sections.

5.3.3 The internal structure of vertical homophony

In the rest of this section, I will scrutinise the structure of the 55 paradigms with a vertical homophony. Different kinds of vertical homophony are attested and it is highly instructive to take a closer look at the frequencies of the various kinds of homophonies. In general, the participants that are involved are an important basis on which to explain why certain kinds of homophony occur more often than others. The relevant parts of the person hierarchy (see section 5.2) are used to account for the differences in frequency. A higher position in the person hierarchy generally leads to a greater frequency. In this section, the 41 cases of a vertical homophony without an inclusive/exclusive opposition will be discussed first. Following that, the 14 ‘exceptional’ cases of a vertical homophony with an inclusive/exclusive opposition are examined. In both cases, the person hierarchy will be of help explaining the attested preferences.

Without an inclusive/exclusive opposition, the paradigm has maximally three different non-singular morphemes: the ‘unified-we’ (ie 1+2, 1+2+3 and 1+3 combined), the second person plural (ie 2+3) and the third person plural (ie 3+3). There are four theoretically possible kinds of vertical homophony with these three elements. These four theoretical possibilities are presented in the first column of Table 5.5. All these four theoretical possibilities are attested; the frequencies of each homophony is shown in the second column. On the basis of these frequencies, roughly three different grades can be distinguished. First, the most commonly found structure is separate marking of 3+3 opposed to a homophony of the other categories. Second, separate marking of 2+3 against the rest and separate marking of ‘unified’ first person complex against the rest are equally frequent. Finally, a complete homophony of all non-singular catego-

ries is least frequent. This frequency-effect is summarised in the hierarchy (5.9), which mirrors part of the person hierarchy.

(5.9) speaker and addressee separate > speaker or addressee separate > none separate

This hierarchy correlates with inflectional marking. As was argued above, all paradigms with a vertical homophony together show a strong tendency to be inflectionally marked (see Table 5.2). Indeed, 75 percent of the 41 cases from Table 5.5 are inflectionally marked. Relative to these general preferences, the stages in the hierarchy (5.9) correlate with the number of inflectional paradigms. The higher in the hierarchy, the lower the proportion of inflectional cases and vice versa. Explicit marking of speaker and addressee (ie high in the hierarchy) shows a preference for independent marking. Less explicit marking (ie lower in the hierarchy) has a preference for inflectional marking. This correlation between explicitness of the paradigm and morphological status will be found time and time again (see especially section 5.8).

Vertical Homophony	Number of cases	Indep.	Infl.	% Inflectional
1+2/3, 2+3	16	6	10	62.5%
1+2/3, 3+3	10	4	15	78.9%
2+3, 3+3	9			
1+2/3, 2+3, 3+3	6	0	6	100%
Total	41	10	31	75.6%

Table 5.5: Kinds of vertical homophony *without* inclusive/exclusive opposition

In contrast to these 41 cases, there are 14 other cases that show a vertical homophony alongside an opposition between the inclusive and the exclusive. These 14 cases are the exceptional cases to the implication ‘vertical homophony → no inclusive/exclusive opposition’. Most of these 14 cases mark the categories 1+2 and 1+2+3 together (ie inclusive) against 1+3 (ie exclusive). Five different kinds of homophony are attested. They are presented in the first column of Table 5.6. There is one case, attested in the Papuan language Kunimaipa, where the category 1+2+3 is marked identical to 1+3 and 2+3, but different from 1+2. This exceptional structure is added at the bottom of Table 5.6. Two kinds of homophony occur clearly more often than the others. The first of these is a homophony between the inclusive and the category 2+3. This cluster of categories can be called the ‘second person complex’ (the complement to the ‘first person complex’). This second person complex is the combination of all non-singular categories that include at least the addressee (ie 1+2, 1+2+3 and 2+3). The other relatively common homophony is the precise opposite of this ‘second person complex’: the ‘non-second person complex’. This homophony combines the exclusive with the third person plural. This cluster includes all non-singular categories that do not include the addressee (ie 1+3 and 3+3). The special status of these two kinds of homophony can be elucidated by referring to the second part of the person hierarchy as described in (5.1) above: speaker > addressee. The speaker has in general greater salience than the addressee. The two relatively common kinds of vertical homophony are both centred around the category ‘addressee’. If the more salient category ‘speaker’ would be taken as the centre, this results in the ‘first person com-

plex. This cluster of categories was taken as the starting point for the classification because it is a characteristic of the large majority of the pronominal paradigms. The relative salience of speaker > addressee is reflected by the much larger number of cases that mark the first person complex by one morpheme (ie all languages that have ‘unified we’) when compared to the few cases that mark the second person complex by one morpheme (ie the five cases noted in the table below). In this sense, the few cases that are marked around the second person are really exceptional cases from a cross-linguistic point of view. Finally, it is tempting to draw conclusions from the fact that the only two independently marked paradigms are found among the relatively more ‘regular’ kinds of homophony. However, the total number of examples is much too low to conclude anything from the distribution of these two cases. The only thing to note is that, across the board, these 14 ‘exceptional’ cases are almost exclusively found among inflectionally marked paradigms.²

Vertical Homophony	Number of cases	Indep.	Infl.	% Inflectional
inclusive and 2+3	5	2	7	77.8%
exclusive and 3+3	4			
inclusive and 3+3	2	0	5	100%
exclusive and 2+3	1			
2+3 and 3+3	1			
1+2+3, 1+3 and 2+3	1			
Total	14	2	12	85.7%

Table 5.6: Kinds of vertical homophony *with* inclusive/exclusive opposition

5.3.4 Summary

Vertical homophony is not a widely acknowledged possibility for pronominal paradigms. It is not a very widespread property, but it is definitively a possibility for human language (55 cases in the total set of 265, which is 20.8%). In this section, the structural characteristics of such paradigms were scrutinised. Two structural constraints were noted on the occurrence of vertical homophony. First, vertical homophony mainly occurs in paradigms that do not have an opposition between inclusive and exclusive in the first person plural. Second, there is a strong tendency for paradigms with a vertical homophony to be inflectionally marked. Next, the internal structure of the different kinds of vertical homophony was discussed. Different parts of the person hierarchy were used to explain the relative frequency of the various forms of vertical homophony. Without an opposition between inclusive and exclusive in the first person plural, the following hierarchy was established:

(5.10) speaker and addressee separate > speaker or addressee separate > none separate

² The vertical homophony between inclusive and third person plural is not a common structure, only attested in two cases in the present sample. Yet, Rodrigues (1990) argues forcefully that in the (extinct) language Tupinambá a parallel is made between ‘You and I’ (ie the inclusive) and ‘Neither You nor I’ (ie the third person plural). Cross-linguistically, this might be an uncommon strategy, but for a particular language it can be completely normal and logical to categorise the world in this way.

When an opposition between inclusive and exclusive is present in the paradigm, the situation is quite different. There seems to be a preference for either the combination of all addressee-based categories into one morpheme, or the combination of all non-addressee-based categories into one morpheme. Finally, it has been observed that paradigms that mark less explicitly the difference between the three basic categories ‘speaker’, ‘addressee’ and ‘other’ in the non-singular, are more often found to be inflectional paradigms. This observation will be elaborated on in section 5.8.

5.4 Singular Homophony

5.4.1 Preamble

Singular homophony is characterised by the fact that the three singular categories (‘speaker’, ‘addressee’ and ‘other’) are only marked by two different morphemes. Two of the three categories are morphologically homophonous. These two function together in paradigmatic opposition to the third category. An example of this phenomenon is presented in (5.11). These suffixes are the Standard Dutch present tense suffixes (without inversion). The categories ‘addressee’ and ‘other’ are marked identical with a suffix *...-t*. This is an example of a singular homophony. In contrast, when all three singular categories are marked by different morphemes, such a paradigm is called ‘split singular’.

(5.11) DUTCH

1	<i>...-∅</i>	<i>...-en</i>	<i>1+2</i>
2	<i>...-t</i>		<i>1+2+3</i>
3	<i>...-t</i>		<i>1+3</i>
			<i>2+3</i>
			<i>3+3</i>

The occurrence of a singular homophony is highly constrained. It occurs only in 25 of the 265 cases in the present sample. This amount is probably even higher than it would be in a balanced sample of the world’s languages, as I have included every attested case with a singular homophony, but left out cases of more widespread paradigmatic structures. However, by including all cases of singular homophony, the number of examples of this exotic phenomenon is raised to a level that allows for a structural analysis.

5.4.2 Some restrictions on singular homophony

Generally, singular homophony is not a preferred structure for a pronominal paradigm. Three different structural constraints on the occurrence of singular homophony will be discussed in this section. First, singular homophony does not occur when there is an inclusive/exclusive distinction; second, it mainly occurs together with a vertical homophony and, finally, it occurs only in inflectional paradigms. As will be argued, these three constraints all lessen the importance of a clearly marked opposition between the participants, opening up to the extreme case of not distinguishing formally between these three singular categories in the paradigmatic structure.

The first observation is that singular homophony is only found with ‘unified-we’ or with ‘no-we’ paradigms. Singular homophony is unheard of when an inclusive/exclusive distinction is attested in the paradigm. The frequencies of the possible combinations are presented in Table 5.7. The distribution of singular homophony can be captured by the implication ‘singular homophony \rightarrow no inclusive/exclusive opposition’. Singular homophony is blocked by an inclusive/exclusive opposition. Or formulated in reverse: singular homophony is made possible by the absence of such an opposition. An explanation can be found in the nature of the inclusive/exclusive opposition. Such an opposition overtly marks the difference between the combination speaker-addressee and speaker-other. Only if this distinction is neutralised, is it possible to conflate singular categories (this insight will be further investigated in section 5.7). It is still not clearly preferred, as only 25 out of 144 cases (17.4%) without an inclusive/exclusive distinction show a singular homophony.

		<i>Inclusive/Exclusive</i>		
		-	+	
<i>Singular Homophony</i>	-	119	121	240
	+	25	0	25
		144	121	265

Table 5.7: Implication ‘singular homophony \rightarrow no inclusive/exclusive opposition’
(Fisher’s exact $p = .000$, $\varphi = -.296$)

Next, the structure of the non-singular marking is an important factor for the occurrence of singular homophony. In Table 5.8, the occurrence of vertical homophony is crossed with the occurrence of singular homophony. It turns out that singular homophony is most commonly found when there is also a vertical homophony in the non-singular. The distribution reveals a strong implication ‘singular homophony \rightarrow vertical homophony’. The explanation for this implication is similar to the one given for the previous restriction. If a vertical homophony is present in the paradigm, this means that the categories 2+3 and 3+3 are homophonous with the first person complex or with each other. In such paradigms, the distinction between the three persons is already neutralised to a large extent in the non-singular, opening up the possibility for a singular homophony. In a paradigm with a vertical homophony, the reference to participants is already mixed up to a great extent, up to the point where it becomes relatively easy for a singular homophony to occur (21 cases out of 55, which is 38.2%).

		<i>Vertical Homophony</i>		
		-	+	
<i>Singular Homophony</i>	-	206	34	240
	+	4	21	25
		210	55	265

Table 5.8: Implication ‘singular homophony \rightarrow vertical homophony’
(Fisher’s exact $p = .000$, $\varphi = .503$)

Finally, the morphological status of the paradigm is an important factor for the occurrence of singular homophony. Singular homophony is only attested in inflectional paradigms. This constraint that was already noted in section 2.4.5 and is quantified here in Table 5.9. The distribution of the 265 cases reveals a strong implication ‘singular homophony \rightarrow inflectional marking’. This constraint can be explained by noting that an opposition that is marked by an inflectional morpheme is cognitively less salient. Of course, inflectional marking has a meaning/function, but it is normally not intuitively clear to the language user (and often neither to the linguist) what is the precise communicational value of an inflectional morpheme. In independent pronouns, the situation is radically different. Independent pronouns are complete words, and words have a great salience to the users of the language. The three categories ‘speaker’, ‘addressee’ and ‘other’ are not to be intermingled in this case. In contrast, in inflectional marking, the possibility exists for a neutralisation of the overt three-way distinction in the singular, although it still is not common (25 cases out of 148, which is 16.9%).

		<i>Inflectional</i>		
		-	+	
<i>Singular Homophony</i>	-	117	123	240
	+	0	25	25
		117	148	265

Table 5.9: Implication ‘singular homophony \rightarrow inflectional marking’
(Fisher’s exact $p = .000$, $\varphi = .287$)

5.4.3 Summary

Singular homophony is a rather unusual phenomenon, although not unheard of in human language. The examples attested are widely dispersed over the world’s languages; they are clearly not areally or genetically conditioned incidents (see chapter 2). Structurally, singular homophony is a highly constrained phenomenon. It mainly occurs when no inclusive/exclusive opposition is present in the paradigm, when there is a vertical homophony in the non-singular and when the paradigm is inflectional. All these constraints open up the possibility for singular homophony, because they cause the paradigm to be less explicitly in the marking of PERSON. The dimension of explicitness of a pronominal paradigm will be further developed in the next sections.

5.5 Explicitness Hierarchy

5.5.1 Preamble

The occurrences of vertical homophony and singular homophony are highly correlated, as was shown in the preceding sections. There is even a strong implication between the two: singular homophony \rightarrow vertical homophony. In fact, this implication is part of a larger hierarchy, which I will call the EXPLICITNESS HIERARCHY. The Explicitness Hierarchy will be formulated as a series of four different kinds of oppositions that can be marked in a pronominal paradigm. Such an opposition will be called

a SPLIT. The definitions of the four different kinds of split that are used in the hierarchy are the following:

- Split inclusive:** there is a special morpheme for the minimal inclusive (ie the category 1+2), differently marked from the augmented inclusive (ie the category 1+2+3);
- Split ‘we’:** there are two different forms for ‘we’: a difference is marked between the inclusive (ie the categories 1+2 and 1+2+3 combined) and the exclusive (ie the category 1+3);
- Split non-singular:** there is *no* vertical homophony, ie the categories 2+3 and 3+3 are marked separately from the first person complex and from each other;
- Split singular:** there is *no* singular homophony, ie the three singular categories 1, 2 and 3 are marked separately.

These four kinds of split are all different ways to discriminate categories of *person*. Starting with the last, a SPLIT SINGULAR means that all three singular persons are marked differently from each other. This is the most basic discrimination of person. The other three kinds of split are different kinds of person discrimination in the non-singular. A SPLIT NON-SINGULAR is, in a way, the mirror-image of a split singular. The ‘plural’ forms of the three singular persons have to be discriminated to be split non-singular. Then there are first two kinds of split, both referring to the first person complex. The first person complex can be split into an inclusive and an exclusive part (SPLIT ‘WE’), and the inclusive part can be split into a minimal and an augmented inclusive (SPLIT INCLUSIVE). A paradigm that has all of these four kinds of splits is maximally explicit as to the marking of the different possible categories of person. A paradigm without any of these splits is minimally explicit in marking person.

5.5.2 The building blocks of the hierarchy

The different kinds of split are correlated with each other in various ways. Many correlations have already been discussed in different places in the present work. A selection of these correlations will be repeated in this section. They will be slightly reformulated to show that the different kinds of split form a hierarchy.

The first correlation to be repeated here was formulated above in section 3.7. There it was argued on qualitative grounds that a paradigm with a split inclusive always has also an exclusive counterpart. This finding is reformulated in (5.12) as an implication. A split inclusive can only occur if there is also a split ‘we’.

(5.12) Split inclusive \rightarrow Split ‘we’

This implication can now also be shown to hold quantitatively in the sample of 265 paradigms. The distribution of the different kinds of split over the different 265 paradigms is presented in Table 5.10. The implication is substantiated by the distribution as shown in this table. The four cases with a split inclusive but without a split ‘we’ are

found in Bardi, Lakhota and Kunimaipa.³ These paradigms were discussed in section 3.6.6.

		<i>Split Inclusive</i>		
		-	+	
<i>Split 'we'</i>	-	144	4	148
	+	90	27	117
		234	31	265

Table 5.10: Implication ‘split inclusive \rightarrow split inclusive/exclusive’
(Fisher’s exact $p = .000$, $\varphi = .315$)

The second correlation to be repeated here is a reformulation of the implication ‘vertical homophony \rightarrow no inclusive/exclusive opposition’ that was first formulated in section 5.3.2. In the terminology as proposed at the start of this section, the absence of an inclusive/exclusive opposition is the equivalent to showing no split ‘we’. A vertical homophony is the same as ‘no split non-singular’. The implication, after it has thus been reformulated, is shown in (5.13). Reversing the direction of the implication removes the negations, resulting in the implication (5.14). For completeness, the distribution as discussed in section 5.3.2 is repeated in Table 5.11 with the new terminology for the two intersecting dimensions.

(5.13) No Split non-singular \rightarrow No Split ‘we’

(5.14) Split ‘we’ \rightarrow Split non-singular

		<i>Split 'we'</i>		
		-	+	
<i>Split Non-Singular</i>	-	41	14	55
	+	103	107	210
		144	121	265

Table 5.11: Implication ‘split we \rightarrow split non-singular’
(Fisher’s exact $p = .001$, $\varphi = .208$)

The last correlation to be repeated here is a reformulation of the implication ‘singular homophony \rightarrow vertical homophony’ that was first formulated in section 5.4.2. Just as in the previous reformulation, vertical homophony is replaced by ‘no split non-singular’. The other part of the implication, singular homophony, is replaced by ‘no split singular’. The implication, after it has thus been reformulated, is shown in (5.15). Reversing the direction of the implication removes the negations, resulting in the implication (5.16). For completeness, the distribution as discussed in section 5.4.2 is repeated in Table 5.12 with the new terminology for the two intersecting dimensions.

³ These four cases are taken together with the 117 cases of split ‘we’ in all other counts that follow. This is not completely along the lines of the definitions presented, but it seems even less appropriate to classify the four exceptional cases with the 144 cases that have no inclusive/exclusive opposition whatsoever.

(5.15) No Split singular → No Split non-singular

(5.16) Split non-singular → Split singular

		<i>Split Non-Singular</i>		
		-	+	
<i>Split Singular</i>	-	21	4	25
	+	34	206	240
		55	210	265

Table 5.12: Implication ‘split non-singular → split singular’
(Fisher’s exact $p = .000$, $\varphi = .503$)

5.5.3 The hierarchy

The three implications (5.12), (5.14) and (5.16) form a hierarchy, as shown by the frequencies in Table 5.13. The five most frequent types, labelled as ‘common’, account for 92.1% of the 265 paradigms. The types that are part of the hierarchy are clearly more common than the other attested possibilities. This hierarchy is called the EXPLICITNESS HIERARCHY because it reflects the order in which different forms of split are expected to occur in a pronominal paradigm: the more referential categories are split (ie marked by different morphemes), the more explicit a paradigm marks referential categories. The most explicit paradigms are the ones that mark all four kinds of split (left on the hierarchy); the least explicit paradigms are the ones that do not adhere to any of the four kinds of split (right on the hierarchy).

	common					semi-common			rare							
<i>Split inclusive</i>	+	-	-	-	-	+	-	-	+	+	+	+	+	-	-	
<i>Split ‘we’</i>	+	+	-	-	-	-	+	-	+	+	-	+	-	-	+	+
<i>Split non-singular</i>	+	+	+	-	-	+	-	+	+	-	+	-	+	-	+	-
<i>Split singular</i>	+	+	+	+	-	+	+	-	-	+	+	-	-	+	-	-
Number of cases	26	78	99	20	21	3	12	4	1	0	0	0	0	1	0	0
	244 (92.1%)					21 (7.9%)										

Table 5.13: The Explicitness Hierarchy

There are five stages in the Explicitness Hierarchy. Note that each stage in the hierarchy does not just stand for one single paradigmatic structure, but that various kinds of paradigms are subsumed in each stage. Selected examples of each stage are presented in Figure 5.3. A few kinds of variation are glossed over in this presentation. First, for all five stages, there are also paradigms with different kinds of horizontal homophony, not shown here (see section 5.6). Second, for stages 4 and 5, there are also other kinds of vertical homophony possible (see section 5.3.3). Finally, for stage 5, there are also other kinds of singular homophony possible (see chapter 2). Notwithstanding this diversity, the examples presented in Figure 5.3 present a clear illustration of the rationale behind the hierarchy.

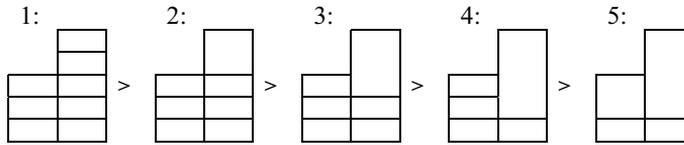


Figure 5.3: Selected examples of the five stages of the Explicitness Hierarchy

The Explicitness Hierarchy starts out with a completely differentiated set of 8 pronominal elements (stage 1). Then more and more non-singular categories are combined into morphemes that are less explicitly marking the referential categories (stages 2, 3 and 4). Only at the very end of the spectrum, it is possible for the singular categories not to be distinguished completely. This comprises the following hierarchy:

(5.17) **Explicitness Hierarchy (rough outline)**

totally explicit > less explicit non-singular > less explicit singular.

The middle part of this hierarchy ('less explicit non-singular') consists of three different stages; the stages 2, 3 and 4 from Figure 5.3 above. The difference between these three stages is governed by the PERSON HIERARCHY. This can be inferred from the verbal description of the five stages Explicitness Hierarchy as presented below.

(5.18) **Explicitness Hierarchy (complete)**

- 1) all categories are different;
- 2) all non-singular categories that include SPEAKER AND ADDRESSEE are marked together;
- 3) all non-singular categories that include AT LEAST THE SPEAKER are marked together;
- 4) different non-singular categories that include THE SPEAKER OR THE ADDRESSEE are marked together;⁴
- 5) some singular categories are marked together on top of that.

The middle part of this hierarchy is based on a special form of the person hierarchy, applied to non-singular categories (see also section 5.2):

(5.19) **Person Hierarchy (non-singular)**

speaker and addressee > at least speaker > speaker or addressee

Looking back at the frequencies of the different types from the Explicitness Hierarchy in Table 5.13, it is clear that relative to each other the five major types are not equally common. The preference for the mid-types in the hierarchy is strengthened by the predominance of horizontal homophony in these types (see especially section 5.6.3). Among the other attested types (the ones not observing the Explicitness Hierarchy in

⁴ This description of stage 4 subsumes many different kinds of paradigms. To be precise, the list should be expanded with all the possibilities of vertical homophony that were discussed in section 5.3.3. These are the following (in order of descending frequency): **A**: all categories which INCLUDE AT LEAST THE SPEAKER OR THE ADDRESSEE are marked together (ie 1+2, 1+2+3, 1+3 and 2+3); **B**: all categories which DO NOT INCLUDE THE SPEAKER are marked together (ie 2+3 and 3+3); **C**: all categories which DO NOT MARK THE ADDRESSEE WITHOUT SPEAKER are marked together (ie 1+2, 1+2+3, 1+3 and 3+3); **D**: ALL NON-SINGULAR CATEGORIES are marked together (ie 1+2, 1+2+3, 1+3, 2+3 and 3+3).

Table 5.13), it is not complete randomness that rules. Three types are attested in more than one case. These are designated as ‘semi-common’ in Table 5.13. These three ‘semi-common’ types also follow the Explicitness Hierarchy, although in a slightly deformed way. If the hierarchy is seen as an ordered sequence of choices, then these three types are the result of a single premature decision. Reading the Explicitness Hierarchy from left to right, the first choice is whether to split the inclusive or not. The 26 cases in the first column decide ‘yes’: the inclusive is split into two different morphemes. As a result, all other decisions down the hierarchy are also fixed. The Explicitness Hierarchy predicts that only if the answer is ‘no’, a further choice comes into the picture. In this way, the choices are linked through the hierarchy. However, the ‘semi-common’ types show that it is possible (though strongly discouraged) to go one choice down the hierarchy if the decision on a choice is ‘yes’. These three types follow the Explicitness Hierarchy, but after the first ‘yes’, there follows one ‘no’, only to be continued with other ‘yes’-decisions. The Explicitness Hierarchy allows for a little tinkering, but only at the boundary of the hierarchy. Just as with the real Explicitness Hierarchy, a preference for the mid-type is attested in this deformed hierarchy. The type halfway through the deformed hierarchy is clearly more common than the other two. This frequency difference is probably an effect of the frequency variation of the Explicitness Hierarchy. The more frequent types from the Explicitness Hierarchy also give rise to more frequent deviations from the hierarchy.

5.5.4 Summary

In the first sections of this chapter, vertical and singular homophony were found to be typologically related to each other. On this basis, it has been shown here that these kinds of homophony are also related to the different possible structures of the first person complex. Together these characteristics of a pronominal paradigm form the Explicitness Hierarchy. The basic claim of this hierarchy is that paradigms first reduce explicitness in the non-singular before the explicitness of the singular is reduced.

(5.20) Explicitness Hierarchy (rough outline)

totally explicit > less explicit non-singular > less explicit singular

The middle part of this hierarchy can be even more strictly formulated. Exactly which explicitness is lost in the non-singular is also highly constrained by the Explicitness Hierarchy. First all categories that include both speaker and addressee are combined into one morpheme. Next, all categories that include at least the speaker are all taken together. Finally, different possible categories that include either speaker or addressee are combined:

(5.21) Explicitness Hierarchy (middle part)

speaker and addressee > at least speaker > speaker or addressee

The different stages of the Explicitness Hierarchy are not concerned with horizontal homophony. Indeed, horizontal homophony is found throughout the various stages of the Explicitness Hierarchy. In the next section, the structure of horizontal homophony is scrutinised.

5.6 Horizontal Homophony

5.6.1 Preamble

Horizontal homophony is characterised by the fact that a paradigm shows an overlap between singular and non-singular categories. One morpheme in the pronominal paradigm is used to mark for both a singular referent and a non-singular referential category. A famous example of a horizontal homophony is the English pronoun *you*.

(5.22) ENGLISH

		we	$1+2$
			$1+2+3$
1	I		$1+3$
2	you		$2+3$
3	he/she/it	they	$3+3$

The name ‘horizontal’ homophony is obviously chosen because these morphemes are graphically represented by a horizontal bar in the diagram. However, the definition of a horizontal homophony allows for more variation. An aberrant case of horizontal homophony is, for example, found in the German inflectional suffix *...-t*, presented in (5.23). In this case, there is no horizontal bar in the diagram, but a diagonal connection between the third person singular and the second person plural. Such cases of ‘diagonal’ homophony are considered to be a subset of horizontal homophony. Diagonal cases will turn out to be rather scarce.

(5.23) GERMAN.

		...-en	$1+2$
			$1+2+3$
1	...-e		$1+3$
2	...-st	...-t	$2+3$
3	...-t	...-en	$3+3$

It is possible for a horizontal homophony to co-occur with a vertical homophony. Such a case is, for example, found in Huave, as shown in (5.24). I have interpreted such cases as a combination of a horizontal homophony (between 3 and 3+3) and a vertical homophony (between the inclusive and the third person).

(5.24) HUAVE

		a-...	$1+2$
			$1+2+3$
1	sa-...		$1+3$
2	i-...		$2+3$
3	a-...		$3+3$

5.6.2 A hierarchy of horizontal homophony

There are four different kinds of horizontal homophony. The first person singular can be homophonous with the inclusive ($1+2$ and $1+2+3$ combined) or with the exclusive

(1+3), or of course with both.⁵ These two kinds of homophony will be represented by the notation '1→Inclusive' and '1→Exclusive' respectively. Next, the second person singular can be homophonous with the second person plural (glossed as 2→2+3) and the third person singular can be homophonous with the third person plural (glossed as 3→3+3). Other combinations of singular and non-singular categories are called 'diagonal'. The various combinations of horizontal homophony that are found in the sample are presented in Table 5.14. There are five combinations that are clearly more frequent than the others (shown on the left of the table). Diagonal cases are rather scarce.

1 → Inclusive	-	-	-	-	+	-	-	+	+	+	-	-	+	+	+	+	diagonal
1 → Exclusive	-	-	-	+	+	-	+	+	+	+	+	+	-	-	-	-	
2 → 2+3	-	-	+	+	+	+	+	+	-	-	-	-	-	-	+	+	
3 → 3+3	-	+	+	+	+	-	-	-	+	-	+	-	+	-	+	-	
Number of cases	136	38	22	18	23	6	3	2	4	2	1	1	1	1	0	0	7
	237 (89.5%)					21 (7.9%)						(2.6%)					

Table 5.14: The Horizontal Homophony Hierarchy

The five most frequent combinations of horizontal homophony form a hierarchy. This hierarchy, presented in (5.25), shows in which order the non-singular categories are marked by horizontal homophony (the brackets are only added for typographic clarity). This hierarchy exactly follows the person hierarchy as presented in section 5.2.⁶

(5.25) Horizontal Homophony Hierarchy

No Homophony < (3→3+3) < (2→2+3) < (1→Exclusive) < (1→Inclusive)

The five types that form the hierarchy are not all equally frequent. No horizontal homophony at all is clearly more common than the other options (136 cases). The next stage of the hierarchy, a homophony 3→3+3, is also more common than the other three stages (38 cases). The remaining three combinations are all roughly equally frequent (22, 18 and 23 cases). This frequency effect follows the leftmost part of the Horizontal Homophony Hierarchy. The resulting frequency-hierarchy can be formulated as follows (the most frequent side is less marked):

(5.26) Frequency of Horizontal Homophony

no homophony < speaker NOR addressee involved < speaker OR addressee involved

⁵ There is only one case in the sample (the perfective suffixes from the Papuan language Kunimaipa, see page 89) that shows a horizontal homophony involving only the category 1+2 only. In all other cases, the two inclusive categories 1+2 and 1+2+3 are given the same treatment as far as horizontal homophony is concerned. For this section, I have classified the exceptional case of Kunimaipa together with the other 'combined' inclusive cases.

⁶ Hierarchies closely resembling the Horizontal Homophony Hierarchy were proposed by Forchheimer (1953:12) and Croft (1990:149). There is an indirect link between these two claims, as Croft based his claim on the data from Ingram (1978) and Ingram's work was based completely on the data from Forchheimer. In between these two, Ingram completely missed this hierarchy, because he only based his analyses on the four major paradigmatic structures that he identified amidst the wide variety of structures discussed by Forchheimer.

The other combinations of horizontal homophony are not distributed randomly. The most frequent cases still follow the upper part of the Horizontal Homophony Hierarchy, but mess up on the lower part. As can be seen in Table 5.14, the relatively most frequent combinations only miss the homophony 3→3+3 (total 11 cases). The next two cases do not follow the hierarchy in both 3→3+3 and 2→2+3 (total 6 cases). These cases are again less frequent.

5.6.3 Horizontal homophony and the Explicitness Hierarchy

It is very well possible for horizontal homophony to occur in different stages of the Explicitness Hierarchy. However, there are strong typological constraints on the occurrence of horizontal homophony. The Explicitness Hierarchy is crossed with the Horizontal Homophony Hierarchy in Table 5.15. The combinations that are most frequent are printed in boldface. As can be seen, the amount of horizontal homophony is highest in the middle of the Explicitness Hierarchy; at both ends of the Explicitness Hierarchy the fraction of horizontal homophony is almost absent. The two middle stages in the Explicitness Hierarchy account for more than 75% of the occurrences of horizontal homophony.⁷

		<i>Explicitness Hierarchy</i>					<i>others</i>
		+	-	-	-	-	
<i>Horizontal Homophony Hierarchy</i>	none	24	38	38	14	14	8
	3→3+3	1	14	19	0	1	3
	2→2+3	0	8	10	0	1	3
	1→Excl	0	15	20	0	3	3
	1→Incl	0	0		0	3	0
	others	1	3	12	6	2	4

Table 5.15: Correlation between the Explicitness Hierarchy and the Horizontal Homophony Hierarchy. Horizontal homophony is most commonly found in the middle range of the Explicitness Hierarchy ($\chi^2 = 51.78$; DF = 20; $p < .0001$).⁸

This curious distribution can be summarised in two strong constraints. First, paradigms with vertical and/or singular homophony disfavour horizontal homophony. Second, paradigms with a division between minimal and augmented inclusive disfavour horizontal homophony. The first of these constraints can be inferred from the lower right part of the table. The two rightmost columns of the Explicitness hierarchy represent the paradigms with a vertical and/or singular homophony. These columns

⁷ Three boxes in the table are bigger than the others, since for the lower part of the Explicitness Hierarchy the opposition between inclusive and exclusive is not relevant.

⁸ For the calculation of the chi-square, the 1→exclusive and 1→inclusive rows are combined into one row. Note that the cases that do not follow either of the two hierarchies have roughly chance distribution.

hardly show any horizontal homophony. It seems as if a pronominal paradigm has to choose from a certain point onward in which way to reduce the explicitness of marking. The paradigm may have horizontal homophony, but then it can not also have vertical and/or singular homophony, and vice versa. The combinations of both kinds of homophony in a paradigm is strongly disfavoured. The second constraint is that a division between minimal and augmented inclusive disfavours horizontal homophony. This can be inferred from the lower left part of the table. The leftmost column of the Explicitness Hierarchy represents the paradigms that are completely explicit in the marking of the non-singular categories. The principal characteristic of these paradigms is the opposition between the minimal inclusive (1+2) and the augmented inclusive (1+2+3). These paradigms hardly show any horizontal homophony. This constraint can be explained by considering the difference between the minimal and the augmented inclusive as the ultimate addition for a pronominal paradigm. This opposition is only found in paradigms that distinguish all other referential categories as well. The curious form of this correlation between the Explicitness Hierarchy and the Horizontal Homophony Hierarchy will be used as a hypothesis to describe the diachronic dynamics of pronominal paradigms. The attested changes of paradigmatic structure will roughly follow the lines of the interrelated hierarchies as sketched in this section. This theme will be extensively dealt with in the next chapter.

5.6.4 Summary

Horizontal homophony is characterised by the fact that a morpheme is used for both singular and non-singular referential categories. The occurrence of horizontal homophony in a pronominal paradigm is highly constrained. Of all the theoretically possible combinations of the singular and the non-singular categories, only 4 are attested with relative frequency. These four are the combinations (1→Inclusive), (1→Exclusive), (2→2+3) and (3→3+3). These four are most of the time attested in combinations following the Horizontal Homophony Hierarchy. This hierarchy parallels the person hierarchy.

(5.27) Horizontal Homophony Hierarchy

No Homophony < (3→3+3) < (2→2+3) < (1→Exclusive) < (1→Inclusive)

This hierarchy describes the kind of horizontal homophony that is found in roughly 90% of the paradigms in the current sample. The hierarchy also governs two frequency effects. First, the paradigms that are described by the left side of the hierarchy are the most frequent. Second, the majority of the remaining 10% that does not follow this hierarchy, behaves differently on the left side of the hierarchy. In other words, the more to the right, the fewer the exceptions. Finally, the Horizontal Homophony Hierarchy turns out to correlate in a peculiar way with the Explicitness Hierarchy. Two constraints were formulated. First, paradigms with vertical and/or singular homophony disfavour horizontal homophony. This can be explained by noting that both vertical and horizontal homophony lessen the explicitness of a pronominal paradigm. Seemingly, these two ways to reduce complexity are preferably not combined. Second, paradigms with a division between minimal and augmented inclusive disfavour horizontal homophony. This constraint can be explained by noting that the minimal/augmented opposition is the last opposition to be included in a pronominal para-

digm. Only after all the other (seven) person categories are distinguished morphologically, this opposition can be added to the others.

5.7 Building a semantic map

A semantic map is a special method to summarise the attested patterns in a typological study (Anderson, 1982; Haspelmath, 1997:59-63; Stassen, 1997:577-610). In a semantic map, meanings which are similar are depicted close to each other so as to convey an impression of the internal structure of a semantic field.

‘We can determine “similarity” of meaning typologically. If two particular meanings are often expressed by the same surface form (across a random sample of languages), then we can assume that the two meanings are “similar” to the human mind. ... From “similarities” it is a short step to maps of grammar/meaning space. We arrange different meanings on a map so that “similar” meanings are close together, non-similar meanings farther apart.’ (Anderson, 1982:227-228)

For the present case, it turns out that it is most helpful to draw two different semantic maps. The similarity between the various categories of pronominal reference changes drastically once there is an opposition between inclusive and exclusive ‘we’ in the paradigm. If there is no inclusive/exclusive opposition, many different kinds of homophony are attested. Under this condition, the semantic map is a closely knit web of categories. However, once an inclusive/exclusive opposition is present in the paradigm, the possibilities for homophony are strongly restricted. Accordingly, the matching semantic map looks rather different.

First, the various kinds of homophony are counted in the subset of 144 paradigms WITHOUT an inclusive/exclusive opposition. The results are presented in Table 5.16. Each number in the table refers to the number of times a particular homophony is attested among the 144 paradigms. A particular paradigm can have more than one homophony, so each number in the table has to be interpreted relative to the total of 144 cases. For example, there are eight cases of a speaker-addressee homophony among the 144 paradigms in the sample. These eight cases are 5.6% of the total of 144 paradigms. A first impression from this table is that almost all possible kinds of homophony are attested and that many kinds of homophony occur in ample quantity.

	2	3	$1+2(+3)$, $1+3$	$2+3$	$3+3$
<i>1</i>	8 (5.6%)	9 (6.3%)	31 (21.5%)	1 (0.7%)	–
2		8 (5.6%)	4 (2.8%)	42 (29.2%)	–
		3	–	2 (1.4%)	58 (40.3%)
			$1+2(+3)$, $1+3$	22 (15.3%)	16 (11.1%)
				$2+3$	15 (10.4%)

Table 5.16: Occurrence of homophony among the 144 paradigms WITHOUT an inclusive/exclusive opposition (the percentages are the proportion relative to the total number of 144 cases)

The kinds of homophony that occur in more than one case are depicted Figure 5.4. This figure takes the form of a semantic map. Categories that are frequently found to

be marked homophonous are shown near each other. Roughly, the farther apart the categories in the figure, the less likely they are to be homophonous. The lines that connect the categories in the figure represent the percentage of cases for each particular homophony. The thickness of the lines is proportional to this percentage. The semantic map is drawn as two concentric circles. The outer circle consists of the singular categories, which are incidentally homophonous (singular homophony). All combinations in the outer circle are equally frequent. The inner circle is formed by the non-singular categories. Homophony between these categories (vertical homophony) is clearly more frequent than homophony between the categories in the outer circle. In the inner circle, the homophony between 'we' and 'you-all' is slightly more frequent than the others. Most frequent are the connections between the inner and the outer circle (horizontal homophony). The frequency rises with the Person Hierarchy $1 < 2 < 3$. Finally, there are a few incidental cases that cross through the cognitive map (diagonal homophony).

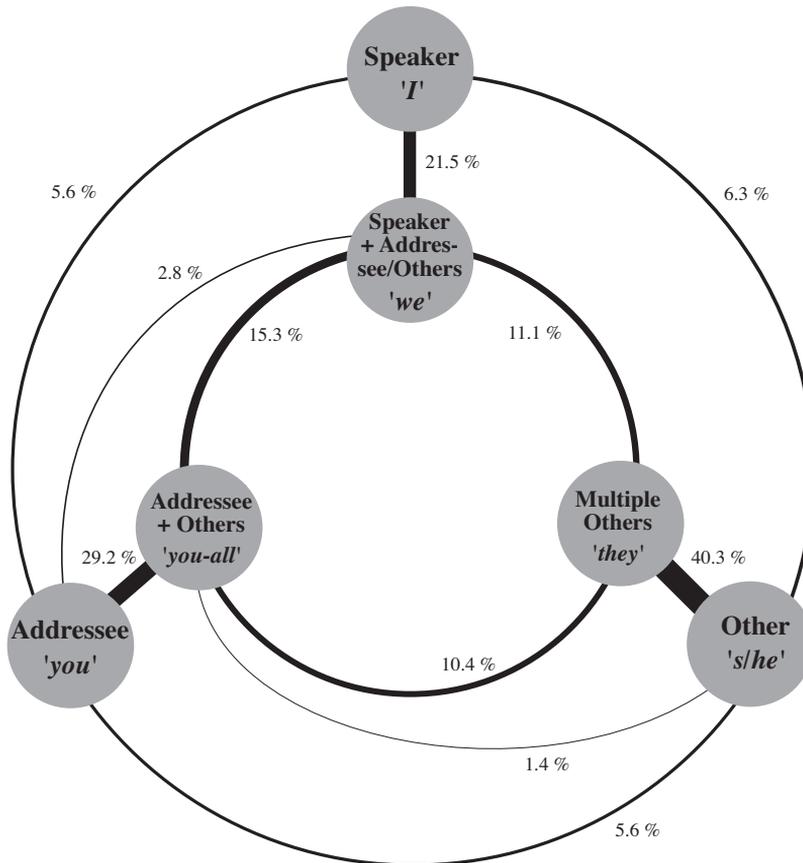


Figure 5.4: Semantic map of paradigms WITHOUT an inclusive/exclusive opposition (the percentages indicate the proportion of each kind of homophony relative to the 144 paradigms without an inclusive/exclusive opposition in the sample)

Next, I turn to the subset of 121 paradigms WITH an opposition between inclusive and exclusive. These paradigms behave rather differently from the previous set as far as homophony is concerned. They show much less vertical homophony and no singular homophony. In contrast, horizontal homophony is common. The occurrence of the various kinds of homophony in these paradigms is shown in Table 5.17. The percentages for each kind of homophony indicate the proportion of occurrence relative to the 121 paradigms with an inclusive/exclusive opposition.

	2	3	$I+2(+3)$	$I+3$	$2+3$	$3+3$
I	–	–	1 (0.8%)	23 (19.0%)	–	–
	2	–	–	–	31 (25.6%)	–
		3	–	–	–	47 (38.8%)
			$I+2(+3)$		5 (4.1%)	1 (0.8%)
				$I+3$	1 (0.8%)	4 (3.3%)
					$2+3$	1 (0.8%)

Table 5.17: Occurrence of homophony among the 121 paradigms WITH an inclusive/exclusive opposition (the percentages are the proportion relative to the total number of 121 cases)

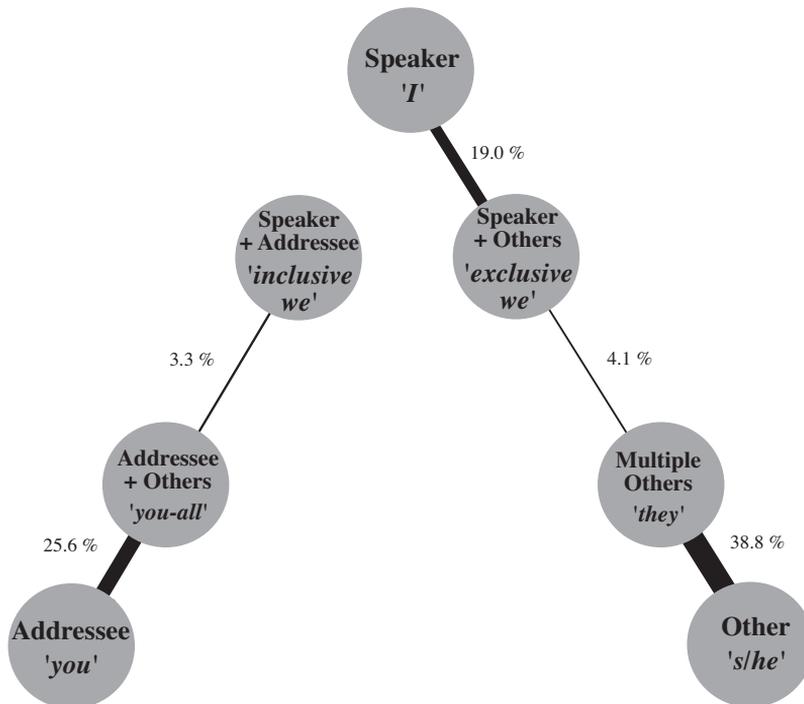


Figure 5.5: Semantic map of paradigms WITH an inclusive/exclusive opposition (the percentages indicate the proportion of each kind of homophony relative to the 121 paradigms with an inclusive/exclusive opposition in the sample)

The kinds of homophony that occur in more than one case are shown in Figure 5.5. This semantic map looks remarkably different from the previous one. The two concentric circles are gone, only the connections between the singular and the non-singular categories have remained (horizontal homophony). The percentages for these connections is roughly identical to the analogous percentages in the previous figure. Except for this correspondence, nothing binds this cognitive map to the previous one. The differences between the two semantic maps are obvious. In this second map, there is no connection between the various singular categories (singular homophony) and only two ‘thin’ connections between the various non-singular categories are attested (vertical homophony). To summarise, once there is an inclusive/exclusive opposition in the paradigm, the possibilities for homophony are strongly restricted.

5.8 Correlations with Morphological Status

5.8.1 Preamble

The present study encompasses the structure of both morphologically independent and inflectionally marked pronominal paradigms.⁹ The morphological status of the person marking was not *a priori* considered to have an influence on the structure of a pronominal paradigm. It is only now, after the structural analysis of the paradigms is completed, that a difference between the two can possibly be shown to exist. In the discussion up till now, there have been a few indications that the morphological status plays a role after all. First, it was discovered that singular homophony can only occur in inflectionally marked paradigms (see sections 2.4.5 and 5.4.2). Second, there is a strong correlation (although not without exceptions) between vertical homophony and inflectional marking (see section 5.3.2). Finally, there is an indication of a (slight) inverse correlation between an inclusive/exclusive opposition and inflectional marking (see also section 5.3.2). In this section, these correlations will be shown to be caused by a general correlation between the morphological status of a paradigm and the two structural hierarchies that were developed in the preceding sections. I will show that both the Explicitness Hierarchy and the Horizontal Homophony Hierarchy are correlated with morphological status. In both cases, the correlation points in the same direction: the more oppositions in a paradigm, the larger the proportion of morphologically independent paradigms in the sample. Towards the other side of both hierarchies (*viz* for paradigms with less oppositions marked), the proportion of inflectional paradigms rises in the sample. Interestingly, the paradigms that do not fit into the hierarchies are even more likely to be inflectionally marked than the least explicit paradigms within the hierarchies. Exceptions tend to occur with greater chance in inflectional paradigms. Throughout the following analysis, it should be kept in mind that the chances for a paradigm to be inflectional in the present sample of 265 cases is a little higher than 50%. The mean percentage of inflectional cases is 55.8% (148 inflectional and 117 independent paradigms).

⁹ The difference between independent and inflectionally marked paradigms is not a clear dichotomy, but more of a continuum. I have not bothered too much with the intermediate ‘clitic’ forms. At face value, these are not very frequently described as such in the grammars, and it would be a study in its own right to develop a useful cross-linguistic definition for clitics. Probably, most paradigms that could be classified as clitics ended up as inflectional in this study.

5.8.2 Explicitness Hierarchy

The Explicitness Hierarchy is repeated below in Table 5.18. The percentages of inflectionally marked cases for each stage of the hierarchy have been added. Two correlations that were found earlier can be found in these percentages. First, the paradigms with a singular homophony were all found to be inflectional, as can be seen in the rightmost column of the hierarchy (100% inflectional). Second, the paradigms with an inclusive/exclusive opposition were found to be slightly more independent, as can be seen in the two leftmost columns of the hierarchy (19.2% and 47.4% inflectional). These two correlations turn out to represent the extremes of a continuous correlation between the Explicitness Hierarchy and the percentage of inflectionally marked paradigms, as is shown in Figure 5.6, below the table.

<i>Split Inclusive</i>	+	–	–	–	–	others
<i>Split ‘we’</i>	+	+	–	–	–	
<i>Split Non-Singular</i>	+	+	+	–	–	
<i>Split Singular</i>	+	+	+	+	–	
Number of cases	26	78	99	20	21	
<i>Independent</i>	21	41	42	10	0	3
<i>Inflectional</i>	5	37	57	10	21	18
% Inflectional	19.2%	47.4%	57.6%	50.0%	100%	85.7%

Table 5.18: Percentage of inflectional paradigms in the Explicitness Hierarchy

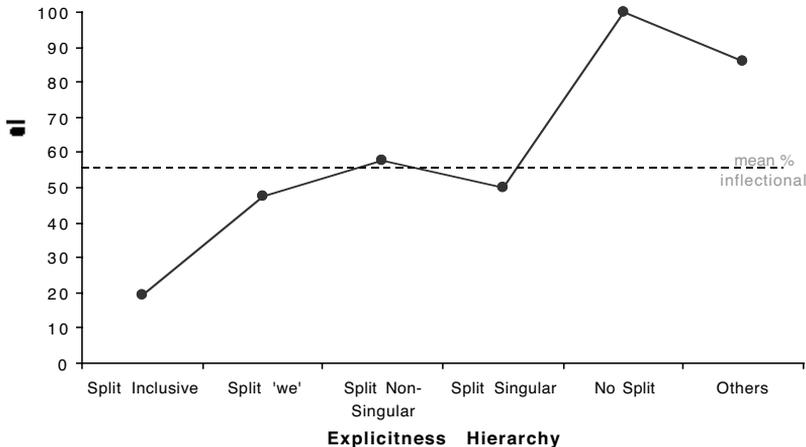


Figure 5.6: Correlation between the Explicitness Hierarchy and proportion of inflectional marking

The graph in Figure 5.6 is not continuously rising. On closer inspection, the ‘dips’ turn out, to be caused by the combination of the various building blocks into one hierarchy. The Explicitness Hierarchy is built from three interrelated implications (see

section 5.5.2). In Table 5.19, the hierarchy is dissected into these three parts, and the percentages of inflectional marking are added to each part individually. The three parts all show a clear correlation with inflectional marking (and the exceptions to the hierarchies are always on the high end). Also, a clear correlation between the three parts is found, as can be seen in Figure 5.7. Each graph subsequently starts and ends higher than the preceding one. The ‘dips’ in the combined graph (see Figure 5.6) occur because most exceptions are found in the middle part of the hierarchy.

<i>Split Inclusive</i>	+	-	-	+	<i>Split 'we'</i>	+	-	-	+
<i>Split 'we'</i>	+	+	-	-	<i>Split Non-Singular</i>	+	+	-	-
Number of cases	27	90	144	4	Number of cases	107	103	41	14
<i>Independent</i>	21	43	52	1	<i>Independent</i>	63	42	10	2
<i>Inflectional</i>	6	47	92	3	<i>Inflectional</i>	44	61	31	12
% Inflectional	22.2	52.2	63.9	75.0	% Inflectional	41.1	55.3	75.6	85.7

<i>Split Non-Singular</i>	+	-	-	+
<i>Split Singular</i>	+	+	-	-
Number of cases	206	34	21	4
<i>Independent</i>	105	12	0	0
<i>Inflectional</i>	101	22	21	4
% Inflectional	49.0	64.7	100	100

Table 5.19: The Explicitness Hierarchy subdivided into its three component hierarchies

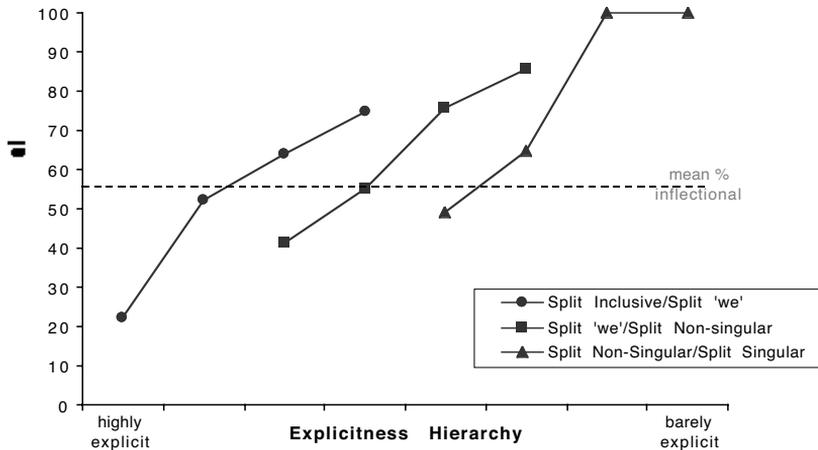


Figure 5.7: Correlation between the subdivided Explicitness Hierarchy and proportion of inflectional marking

5.8.3 Horizontal Homophony Hierarchy

The Horizontal Homophony Hierarchy is also analysed for the proportion of inflectionally marked paradigms. The results are shown in Table 5.20. The percentages are plotted in Figure 5.8. Again, the graph shows a correlation between less explicitness in the paradigm (viz more horizontal homophony) and more inflectionally marked pronominal paradigms in the sample. The correlation is less strong than the one that was found with the Explicitness Hierarchy, but a trend is visible. The paradigms that do not fit into the hierarchy are once again found to be peripheral, with a high proportion of inflectional marking.

<i>1 → Inclusive</i>	–	–	–	–	+	others
<i>1 → Exclusive</i>	–	–	–	+	+	
<i>2 → 2+3</i>	–	–	+	+	+	
<i>3 → 3+3</i>	–	+	+	+	+	
Number of cases	136	38	22	18	23	28
<i>Independent</i>	73	16	7	5	10	6
<i>Inflectional</i>	63	22	15	13	13	22
% Inflectional	46.3%	57.9%	68.2%	72.2%	56.5%	78.6%

Table 5.20: Percentage inflectional marking in the Horizontal Homophony Hierarchy

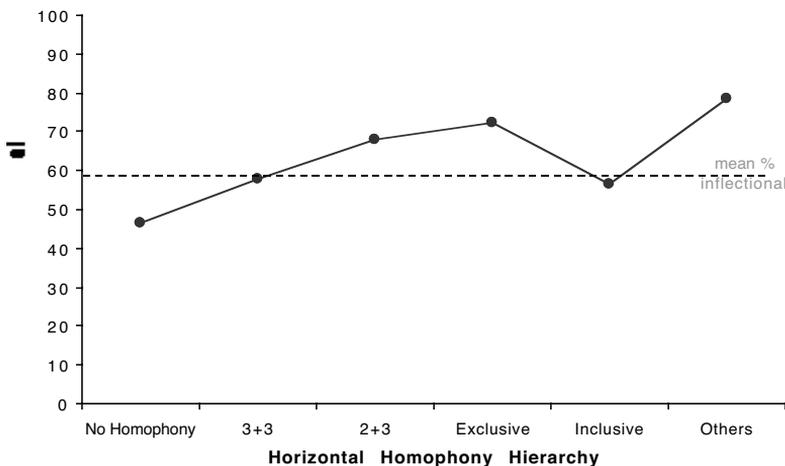


Figure 5.8: Correlation between the Horizontal Homophony Hierarchy and proportion of inflectional marking

The only point on the graph that is slightly off, represents the last stage of the hierarchy. These paradigms are slightly more often independent than expected. The 10 independent paradigms in this stage are the paradigms of the ‘Salt-Yui-type’ (see section 4.4.2). These paradigms are independent pronouns that have no plural pronouns at all. In some cases the singular pronouns can be used to mark for non-singular refer-

ence, in other cases this seems to be impossible. For these examples, the concept of horizontal homophony is perhaps not applicable. Some paradigms simply have no possibility at all to mark non-singular reference. Other linguistic constructions have to be used for that purpose (eg conjunction, compounding, a comitative or an associative construction). These paradigms might be out of place in this category, but because I do not see any better place for them, the ‘dip’ in the graph has remained.

5.8.4 Amount of morphemes

The analysis of the Explicitness Hierarchy and the analysis of the Horizontal Homophony Hierarchy point in the same direction. The more opposition a paradigm has, the likelier it is that the paradigm is morphologically independent. This can also be shown to hold for the present sample if the paradigms are analysed to the number of different morphemes in each paradigm. Exactly what is marked by the different morphemes is left aside, only the number of different morphemes is taken into account here. The number of cases for each amount of morphemes is presented in Table 5.21. The percentage of inflectionally marked paradigms is shown in Figure 5.9 below. As expected, there is a clear correlation between fewer morphemes and more inflectionally marked paradigms.

<i>Amount of morphemes</i>	8	7	6	5	4	3	2
Number of cases	22	44	57	53	58	26	5
<i>Independent</i>	18	29	26	23	11	10	0
<i>Inflectional</i>	4	15	31	30	47	16	5
% Inflectional	18.2%	34.1%	54.4%	56.6%	81.0%	61.5%	100%

Table 5.21: Percentage inflectional marking for amount of morphemes in a paradigm

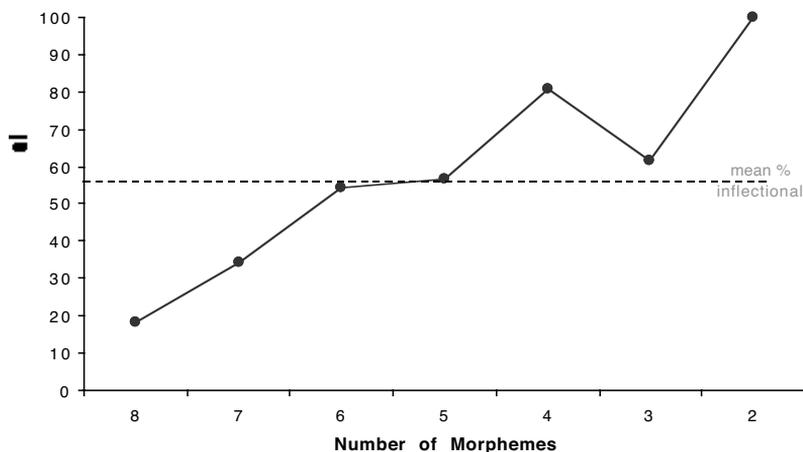


Figure 5.9: Correlation between amount of morphemes and proportion of inflectional marking

This correlation only holds within the boundaries of the present sample. Gender oppositions are not counted; honorific forms are ignored; number marking (dual, paucal etc.) is left out; et cetera. Only the person oppositions in the singular and non-singular were included for the counts, resulting in a maximum range from 2 to 8 forms. Also, the correlations are not absolute; there are very extensive pronominal paradigms that are inflectionally marked and very small paradigms that are independent pronouns. Notably, the inflectional paradigms in the Iroquoian languages in North America and the Gunwinguan language in Australia are among the largest pronominal paradigms to be found in the world. Still, as far as person oppositions are concerned, the correlations clearly show that paradigms with more person oppositions tend to be more often independently marked, and vice versa.

5.8.5 Explanations

These correlations are in need of an explanation. All three correlations point in the same direction: the more oppositions, the higher the proportion of independently marked pronominal paradigms and vice versa. These correlations can be made sense of by two complementary explanations. The first explanation is a cognitive one, invoking a difference between inflectional and independent marking in the knowledge of the language user. The second explanation is structural in nature, invoking different possibilities for inflectional and independent pronominal marking in linguistic change.

First, independent pronouns are generally more salient to the language user (be it speaker or addressee) than inflectional person markers. An independent pronoun is consciously known to the linguistically naive speaker to have a clear referential value. Of course, language users also know the precise referential usage of inflectional elements, but this knowledge tends to be implicit, rather than explicit. Independent words, among them independent pronouns, are known to language users to have a clear ‘meaning’. This ‘meaning’ will be kept referentially clear as much as possible. This urge for referential clarity results in a tendency to have more oppositions in independent pronouns. Also, when a paradigm is not completely explicit, it will normally follow one of the two structural hierarchies (the Explicitness Hierarchy or the Horizontal Homophony Hierarchy), which are both governed by the Person Hierarchy. As a result, the exceptions to the hierarchies always show a high proportion of inflectional marking. Second, an inflectional person paradigm is rather fixed in its structure. It seems to be easier for inflectional person paradigms to lose oppositions than to gain them. It is very well possible for an inflectional paradigm to lose some oppositions, for example by a historical merger. However, to insert extra morphemes in an already existing inflectional paradigm is much more difficult. For independent personal pronouns, I cannot see a comparable structural reason for a difference between losing or adding forms. This diachronic asymmetry turns up in a typological survey as a tendency for inflectional pronominal paradigms to have fewer oppositions.¹⁰

¹⁰ Pending the validity of these two explanations, it has to be noted that the current sample cross-sects through the different possible functional roles of the paradigmatic structure, like subject, object, ergative et cetera. I expect that the effect of any functional difference between the various cases is evenly distributed over the sample. However, future research has to bear out whether this intuition is correct or not.

5.9 Conclusion

In this chapter, a sample of pronominal paradigms has been analysed from the perspective of the individual categories in the paradigms. The reference point for the analyses in this chapter is the ‘maximally explicit’ paradigm, as repeated here in Figure 5.10. Most paradigms do not distinguish all categories by different morphemes. Strong restrictions are attested in the present sample as to which categories are combined together into the reference of one morpheme, and which are separated by different morphemes.

		<i>‘non-singular’</i>			
		1+2	<i>minimal inclusive</i>	} <i>inclusive</i>	} <i>first person complex</i>
		1+2+3	<i>augmented inclusive</i>		
<i>speaker</i>	1	1+3	<i>exclusive</i>		
<i>addressee</i>	2	2+3	<i>second person plural</i>		
<i>other</i>	3	3+3	<i>third person plural</i>		

Figure 5.10: Maximally explicit pronominal paradigm

Different kinds of homophony are attested between the eight categories from the maximally explicit paradigm. Particularly, three different kinds of homophony were distinguished at the start of this chapter (see section 5.2). First, when the three singular categories are not differentiated, this is called SINGULAR HOMOPHONY; in contrast, when the three singular categories are differently marked, this is called SPLIT SINGULAR (see section 5.4). Second, when the second and third person plural are not differentiated from the first person complex, this is called VERTICAL HOMOPHONY; in contrast, when they are differently marked, this is called SPLIT NON-SINGULAR (see section 5.3). Finally, when there is an overlap between singular and non-singular, this is called HORIZONTAL HOMOPHONY (see section 5.6).

The analysis of the various forms of homophony resulted in two hierarchies. Singular homophony, vertical homophony and the possible divisions of the first person complex make up the EXPLICITNESS HIERARCHY (see section 5.5). The various kinds of horizontal homophony form the HORIZONTAL HOMOPHONY HIERARCHY (see section 5.6). Both hierarchies were shown to follow the Person Hierarchy, as repeated below in (5.28). The middle part of the Explicitness Hierarchy showed a slight variant on this hierarchy (see section 5.5.3). The Horizontal Homophony Hierarchy exactly followed the Person Hierarchy (see section 5.6.2). Also, the frequency of the various kinds of vertical homophony was ordered along the lines of the Person Hierarchy (see section 5.3.3). This Person Hierarchy represents a sort of cross-linguistic majority view of the cognitive order in the speech-act situation. A special kind of paradigmatic structure occurs once the highest stage of the Person Hierarchy is reached. If there is an opposition in the paradigm between ‘speaker and addressee included’ and ‘speaker included’ (normally this opposition is referred to as ‘inclusive versus exclusive’), then the variability of the paradigm is strongly constrained (see section 5.7).

(5.28) **Person Hierarchy**

speaker and addressee > speaker > addressee > neither speaker not addressee

Both the Explicitness Hierarchy and the Horizontal Homophony Hierarchy can be summarised as different ways that lead from a minimally to a maximally explicit pronominal paradigm; one path follows through the Explicitness Hierarchy; the other goes through the Horizontal Homophony Hierarchy. Before I present both pathways, two clarifications are necessary. First, there is no indication that these paths are unidirectional. The perspective that is taken here (with a direction from minimal to maximal explicitness) is only a metaphor chosen for convenience to summarise the attested restrictions on the cross-linguistic variation. Second, the two paths are not independent from each other. Notably, the Horizontal Homophony Hierarchy only operates in the middle part of the Explicitness Hierarchy (see section 5.6.3).

The first path from minimal to maximal explicitness leads through the Explicitness Hierarchy (see especially section 5.5.3). This path is illustrated in Figure 5.11 by a few selected paradigmatic structures. It starts out with paradigms with a distinction between singular and non-singular, but the person distinctions are blurred. Homophony is found both in the singular and in the non-singular. Proceeding to the next stage, the singular categories are the first to be separated. Then, once the singular homophony is resolved and the three singular categories are distinctively marked, the non-singular (= ‘vertical’) homophony is removed. At this point, a paradigm is found with separate second and third person plural, but with a ‘unified’ first person complex. Next, the inclusive/exclusive opposition in the first person complex is introduced and, finally, the distinction between a minimal and an augmented inclusive is added, leading to the maximally explicit paradigm.

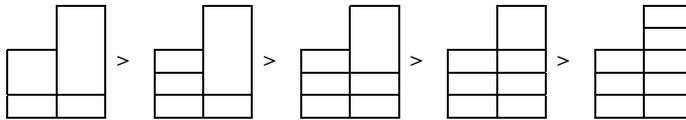


Figure 5.11: Path to maximal explicitness through Explicitness Hierarchy (shown are a few illustrative paradigmatic structures, not all possibilities)

The second path from minimal to maximal explicitness leads through the Horizontal Homophony Hierarchy (see especially section 5.6.2). This path is illustrated by a few selected paradigmatic structures in Figure 5.12.

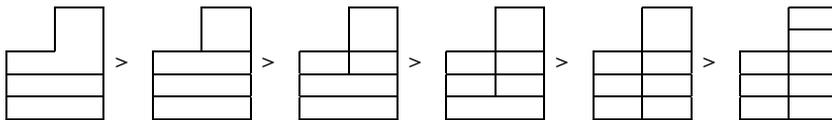


Figure 5.12: Path to maximal explicitness through Horizontal Homophony Hierarchy (shown are a few illustrative paradigmatic structures, not all possibilities)

At the outset of this path, there is no separation between singular and non-singular. There is horizontal homophony throughout the paradigm. The first non-singular category to be separated from the singular is the inclusive, then follows the exclusive. Next comes the second person plural and, finally, the third person plural completes

the separation. At this point, there is no overlap between singular and non-singular anymore. The minimal inclusive can only now be separated from the augmented inclusive.

Both paths correlate with the morphological status of the marking. At the minimally explicit side, a preference for inflectional marking is found. In contrast, a preference for morphologically independent marking is attested at the maximally explicit side. This correlation is explained by two different corroborating tendencies. First, language users are consciously using independent pronouns, whereas inflectional marking is much less salient to the naïve language user. This leads to more explicitness in independent pronouns. On the other hand, an inflectional paradigm can be more easily reduced (through historical merger) than grow in size (through addition of new inflectional categories in a paradigm). This asymmetry is not applicable to independent pronouns. This diachronic difference between inflectional and independently marked pronouns leads to a predominance of inflectional paradigms at the less explicitly marked side of the spectrum (see section 5.8). Both explanations can also shed light on the observation that ‘irregular’ paradigms have a strong tendency to be inflectional. With ‘irregular’ paradigms, I refer to those paradigms that do not fit into either of the two hierarchies. The referential structure of such paradigms is strangely organised, something which speakers mainly allow to develop by merger in inflectional paradigms.

Towards the end of this chapter, more and more diachronic metaphors were used to describe the typological patterns that have been found. In all cases, these descriptions should be read as metaphors and not as real diachronic claims. The analyses presented have been purely synchronic, so that no diachronic conclusions are permitted. In the next chapter, I will develop a method with which it is possible to peer into the diachronic dimension.

Chapter 6

Connecting paradigms

Person paradigms through time and space

6.1 Introduction

Pronominal elements are among the most popular linguistic items for comparative research into the history and prehistory of language. For example the famous *Analytical Comparison of the Sanskrit, Greek, Latin and Teutonic Languages* by Bopp (1820), one of the major early landmarks of the historical-comparative method, is a comparison of pronominal elements in different Indo-European languages. Pronominal elements have remained an important source of information in the historical-comparative line of research until the present day. In this extensive body of work, the locus of investigation has always been the individual pronominal elements. Although it is widely acknowledged that pronominal elements are bound into a paradigm, the paradigmatic structure as a whole has never been consistently included in a diachronic investigation. Such an investigation would ideally lead to a history of the structure of the paradigm, which is a different level of analysis altogether.

In this chapter, an attempt will be made to formulate a first outline of a theory on change of pronominal paradigms as a whole. This outline will be based on the typological investigations from the previous chapter. In the previous chapter, the major restrictions of the occurrence of particular oppositions were summarised in two hierarchies: the Explicitness Hierarchy and the Horizontal Homophony Hierarchy. In this chapter, these synchronic hierarchies will be taken as hypotheses for the options in the change of paradigmatic structure. In section 6.2, I present the method that will be used in this chapter for testing whether typological restrictions as formulated in the previous chapter can be interpreted diachronically. The basic principle is to compile examples of cognate paradigms, ie pronominal paradigms from closely related languages that show only small differences. Such small differences present a window on possible historical changes. In the next two sections, the two hierarchies that were formulated in the previous chapter are evaluated on the basis of attested cognate paradigms. First, in section 6.3, the Horizontal Homophony Hierarchy is scrutinised. This hierarchy turns out to be only partly interpretable in diachronic terms; paradigms seem to be able to move with big strides through this hierarchy. It is not necessary for a paradigmatic change to follow all separate stages of this hierarchy. However, it is not all chaos that rules; there remain some barriers that paradigms do not seem to cross directly when moving through time and space. Second, in section 6.4, the Explicitness

Hierarchy is tested on the diachronic interpretation. This hierarchy turns out to stand the test rather well. At least, this hierarchy remains a good candidate for further diachronic investigation. Finally, in section 6.5, the remaining diachronic connections from the two hierarchies will be combined into a cognitive map of interconnected paradigmatic structures.

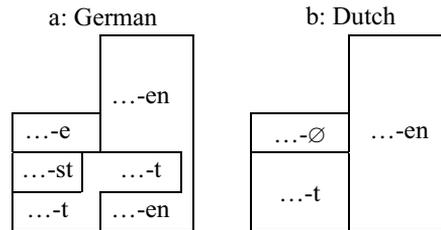
6.2 Cognate paradigms

The method that will be used to approach the diachronic dimension is to compile examples of COGNATE PARADIGMS and then interpret such exemplars as a window on the interrelations among different paradigmatic structures. The basic idea behind a set of cognate paradigms is that the paradigms are highly similar, but have a slightly different paradigmatic structure. Such cognate paradigms will be interpreted as showing a small (and perhaps even the smallest possible) paradigmatic change. In most cases that will be discussed, no stance will be taken as to the direction of the change. It would entail much more in-depth analysis of the individual cases to establish a direction and would require too much analytic work to be able to retain the world-wide (and thus a somewhat coarse-grained) perspective that is taken in the present work. Only a first step towards unravelling the diachronic dynamics of paradigmatic structure is taken in this chapter. This first step is to lay out the general lines of variation in the form of a cognitive map of interrelated paradigmatic structures. I hypothesise that paradigmatic change will proceed along the lines of this cognitive map.

For two pronominal paradigms to be branded as ‘cognate’, a few criteria have to be met. First, the two paradigms have to be found in two closely related languages. This criterion ensures that the similarity between the various paradigms is not a chance phenomenon. The relatedness between the languages has to be established on independent grounds, not on the basis of the form of the pronominal elements alone. For example, the Trans New Guinea Phylum (a well-known problematic case for genetic classification in itself) is to a large extent established on the basis of resemblance between pronominal elements. This basis is far too small to take the next step and compare the form of the pronominal paradigms. In general, the cognate paradigms that are presented in this chapter are found in very closely related languages, often on the verge of being dialectal variants. In most cases, the close relatives that will be taken into the comparison are both extant languages that are geographically not too distant from each other. However, in some of the cases that will be discussed in this chapter, the cognate paradigms come from ‘diachronic’ close relatives, ie from variants of the same language from different points in time. It is of lesser importance to the current aim whether the earlier variant is a direct precursor of the later variant or whether it is a separate though extinct branch. For example, Gothic and present day German are such diachronic close relatives, although Gothic is not a direct precursor of present day German. The method used here is not a historical comparison, but a typological comparison that starts from the broad typological generalisation and tunes in on the fine-grained differences within a genetic group. It is necessary, but not sufficient for the languages to be cognate. A second criterion for pronominal paradigms to be called ‘cognate’ is that the paradigms at hand have to be cognate themselves. This means that the individual elements in the paradigm are related both in form and func-

tion and that the syntagmatic role of the paradigm is roughly identical within the structure of both languages. Consider as an example the present tense suffixes from Standard German and Standard Dutch (without inversion), as shown in (6.1). There is no doubt about the close relationship between the two languages within the West Germanic branch of the Indo-European stock. More importantly, the suffixes themselves show a strong resemblance; this resemblance is seen most strongly in the suffixes *...-t* and *...-en*, which are related both in form and function. Also, both paradigms play a roughly comparable role in the structure of the language as both are used to mark present tense by suffixation to a verb. A comparison between the German inflection and the Dutch independent pronouns would not be a valid case of cognate paradigms because the role of these paradigms in the language is utterly different.¹

(6.1) GERMANIC



The final criterion for two paradigms to be called ‘cognate’ is that the paradigms have to show small paradigmatic differences. These paradigmatic differences are the crux of the present chapter, as they present a window on the dynamics of paradigmatic structure. The preceding criteria ensure that the compared paradigms are almost identical. The attested differences in paradigmatic structure are thus small (and possible the smallest possible) differences in paradigmatic structure. For example, the singular homophony in Dutch present suffixes (addressee and other are marked identical by *...-t*) is not found in German.

To summarise, four criteria have to be met. Cognate paradigms have to be found in cognate languages, and the great majority of the individual morphemes in the paradigms have to be cognates themselves. Also, the paradigms have to have a functionally and formally comparable status within the language as a whole. Finally, the paradigms have to show some differences in their paradigmatic structure. This whole set of criteria is designed to leave only those cases that unambiguously show differences in paradigmatic structure that are small. These small differences can probably arise easily. More precisely, the small differences in paradigmatic structure may arise quickly when compared to phonological changes. The differences between cognate paradigms may arise before the phonological changes have made it difficult (or even impossible) to establish a relation between the paradigms. On the (somewhat precarious) assumption that phonological change is roughly a constant factor in language change, the paradigmatic differences between cognate paradigms represent minimal changes. However, I will not argue for any diachronic interpretation of the cognate

¹ Of course, in this case, the German inflection and the Dutch independent pronouns are also completely dissimilar in form, which already disqualifies them as cognate paradigms on the morphophonological criterion alone. This is not necessarily so.

paradigms.² Examples of cognate paradigms are used to argue that certain paradigmatic structures are similar. Paradigmatic structures can be shown to be closely connected by the existence of a set of cognate paradigms. All the examples of cognate paradigms that will be discussed in this chapter result in a web of interrelated paradigmatic structures; some more closely related than others.

A major problem with this method is that it is difficult to deal with connections that are not attested as a set of cognate paradigms. It is very well possible that, given a more extensive search, other examples would turn up eventually. In incidental cases, it is possible to construct an argument for the implausibility of an unattested connection by analysing the areal distribution of the various patterns. If two different paradigmatic structures are never found in the same geographical area, this is taken as an indication that the two structures are not closely connected (see especially section 6.4.2 below). In most cases, however, the method of cognate paradigms can only be used as a way to test the validity of an independently formulated claim of paradigmatic connections. Given a hypothesis about how pronominal paradigms might be interrelated, examples of cognate paradigms can substantiate such a claim. It is in this sense that the examples in the present chapter should be interpreted.

‘Synchronic typologies function merely as heuristic, though often indispensable, devices in defining the problems and in assembling the relevant data.’ (Greenberg, 1969:194)

The hypothesis to be tested in this chapter is that pronominal paradigms change along the lines of the two hierarchies that were formulated in the previous chapter: the Explicitness Hierarchy and the Horizontal Homophony Hierarchy. These two hierarchies were formulated independently, but turned out to be correlated (see section 5.6.3). The intersection of the two hierarchies showed a clear preference for particular paradigmatic structures. These preferred structures are shown in Figure 6.1, ordered along the lines of the two hierarchies. The Explicitness Hierarchy is shown horizontally and the Horizontal Homophony Hierarchy is shown vertically. The arrows between the paradigmatic structures represent the hypothesis about the expected cognitive map of paradigmatic structure. If paradigmatic change conforms to the lines of the hierarchies, then examples of cognate paradigms are expected to turn up as links between adjacent paradigmatic structures in the cognitive map. These examples only indicate that the hypothesis is feasible; they should not be read as an argumentation for a universal theory of paradigmatic change. Far from that, the possibilities for change are numerous and the examples that will follow shortly are only the first step to a encyclopedic collection of attested cases, which hopefully will lead in the future to a fully-fledged theory of possible changes of a pronominal paradigm. Examples of cognate paradigms form a window on a small part of the cognitive map, and a large set of such windows will hopefully give a better view on the outline of the whole landscape of paradigmatic variation. It is this prospect that drives the search for the minutiae of paradigmatic variation.

² The different paradigms in a set of cognate paradigms are probably derived from some common ancestor. This ancestor could be very similar to one of the extant variants, in which case one of the variants could be said to be historically derived from the other. Otherwise, both variants go back to a further – different – paradigmatic structure at the proto-stage.

The search for cognate paradigms has been restricted to the kinds of paradigm shown in Figure 6.1. These paradigmatic structures represent the nine common cases and the six semi-common cases summarised in section 4.8. Together, these 15 paradigmatic structures account for 78% of the attested variation.³ The remaining 22% ‘rare’ cases do not fit into either of the two hierarchies that will be tested in this chapter. They will be disregarded in this chapter. This restriction to the more widely occurring paradigmatic structures has been made in order to have a chance to find multiple cases of a particular connection between two paradigmatic structures. Good examples of cognate paradigms are not easy to find, which makes it necessary to claim some flexibility if a sufficient number of cases is to be produced.

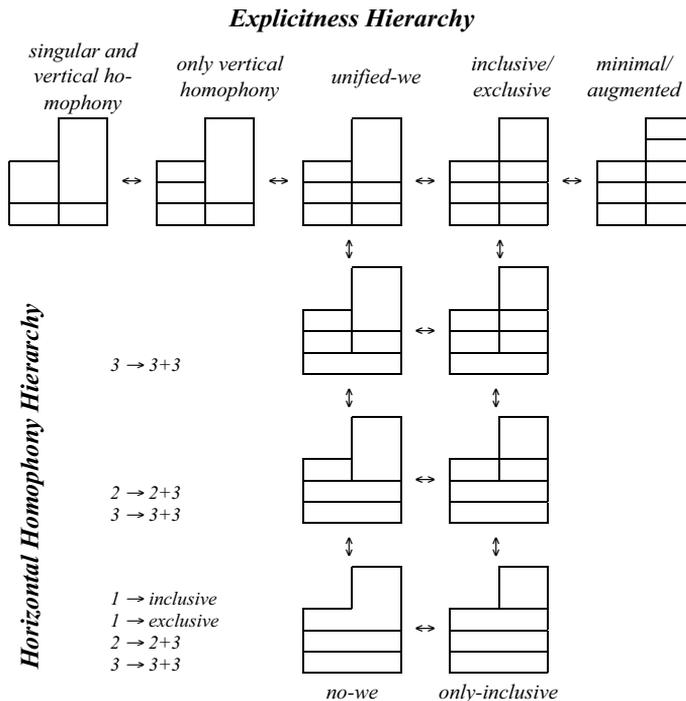


Figure 6.1: Hypothesised cognitive map of the major paradigmatic structures ordered by the Explicitness Hierarchy and the Horizontal Homophony Hierarchy

6.3 Up and down the horizontal homophony hierarchy

6.3.1 Preamble

Most cases of horizontal homophony are attested among the paradigms with ‘unified-we’ (eg the English pronoun ‘we’) and the paradigms with an inclusive/exclusive op-

³ The six semi-common paradigmatic structures are subsumed under the categories ‘singular and vertical homophony’ and ‘only vertical homophony’. The nine common paradigmatic structures are all represented individually in the figure.

position (see Figure 6.1). The various paradigmatic structures on the Horizontal Homophony Hierarchy are repeated in Figure 6.2. The labels that are added to the paradigms in the figure refer to the marking of the first person complex, as discussed in section 3.6. As a hypothesis, I propose that the paradigmatic structure will develop diachronically along the lines of the hierarchy, which would give rise to changes as indicated by the arrows in Figure 6.2. If this hypothesis is any good, it is to be expected that there are pronominal paradigms in closely related languages that only differ in one morpheme on this hierarchy. In this section, I will review a large set of examples attested to substantiate this hypothesis.

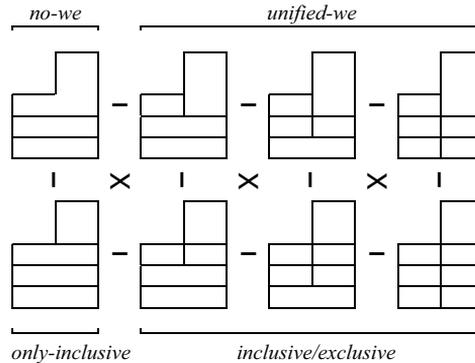


Figure 6.2: The Horizontal Homophony Hierarchy

The connections attested will be described in two parts. First, in section 6.3.2, the connections between the ‘small’ paradigms in the upper left corner of Figure 6.2 are presented. These paradigms (‘small’, as they have only few oppositions) seem to be tightly interwoven, because numerous examples of cognate paradigms are attested between the various paradigms. Second, in section 6.3.3, the connections between these ‘small’ paradigms and the larger paradigms, up the horizontal homophony, hierarchy are presented. The picture becomes somewhat blurred here, indicating that the hypothesis can only be an approximation of the diachronic dynamics of paradigmatic structure. An improved hypothesis is proposed in the summarising section 6.3.4.

6.3.2 Interconnecting the small paradigms

The minimally marked pronominal paradigms are found at the far end of the Horizontal Homophony Hierarchy. These paradigms have horizontal homophony in (almost) all persons; the only specialised non-singular forms are found in the first person complex. In this section, examples of connections between the three smallest paradigms (as highlighted in Figure 6.3) will be discussed.

Connections between these three paradigms are relatively easy to find. Examples from Siouan, Chimbu and Waris are presented to illustrate the connection between these three paradigmatic structures. Some more examples will be shown *en passant* in the next section. In that section, the examples from Arawakan and Macro-Gé are of special interest, as they show that the diagonal connection in this triple is also attested.

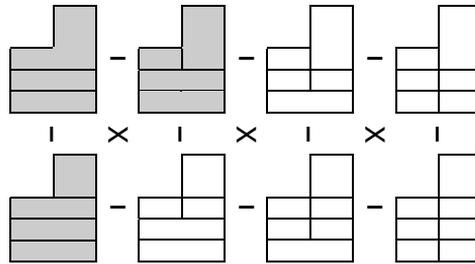
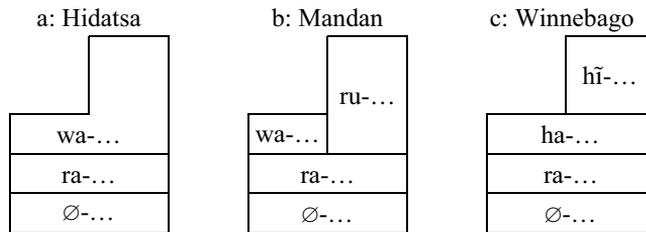


Figure 6.3: The small paradigms of the Horizontal Homophony Hierarchy

Siouan prefixes

The first set of cognate paradigms to be reviewed here is found among the Siouan languages, spoken in northern USA. The agent prefixes of the different Siouan languages show only slightly different paradigmatic structures. This set of cognate paradigms presents an example of the connection between the three small paradigms.

(6.2) SIOUAN

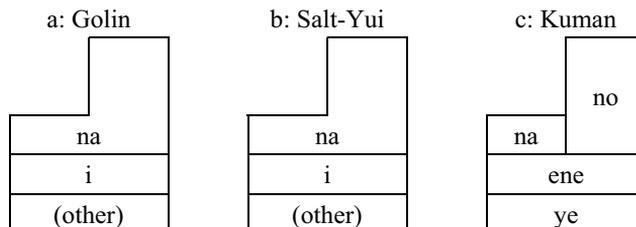


The agent prefixes from Hidatsa are shown in (6.2a). There are no specialised non-singular morphemes in this paradigm, although number can be marked by suffixes (Robinett, 1955: 177; Matthews, 1965: 55, 71). A specialised morpheme for ‘unified-we’ is found in the agent prefixes from Mandan, shown in (6.2b). Otherwise, this paradigm from Mandan is completely identical to the paradigm from Hidatsa (Mixco, 1997: 8). Finally, the agent prefixes from Winnebago are shown in (6.2c). This time, there is a specialised morpheme for inclusive-‘we’ only (Greenberg, 1988: 4-5, citing Susman 1943).

Chimbu pronouns

The Chimbu family is part of the Border stock, one of the parts that constitute the Trans-New-Guinea Phylum in Papua New Guinea.

(6.3) CHIMBU

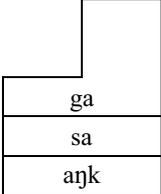
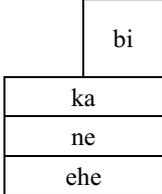
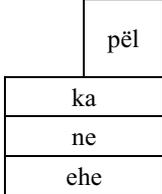


The pronouns from the Chimbu family present an example of the smallest kind of paradigm that exists among the world's independent pronouns. These paradigms further illustrate the link between two different small paradigms. The independent pronouns from Golin (Foley, 1986:70, citing Bunn 1974) and from Salt-Yui (Irwin, 1974:74) are identical. Both paradigms, shown in (6.3 a, b), fail to distinguish any specialised non-singular forms. The independent pronouns from Kuman are slightly different, as can be seen in (6.3 c). Most salient for the present discussion is the existence of a specialised form for 'unified-we' (Foley, 1986:70, citing Piau 1985).

Waris pronouns

The Waris family belongs to the East New Guinea Highlands stock, which is also a part of the Trans-New-Guinea Phylum in Papua New Guinea. The paradigms presented in (6.4) are the short forms of the independent pronouns. These short forms are normally used in fluent speech. The long versions (with more oppositions) are used only when it is necessary to be referentially more explicit.

(6.4) WARIS

a: Manem	b: Amanab	c: Imonda
		

In these paradigms, the phonological correspondences between Manem, shown in (6.4a) and the other two languages are rather opaque; only the first person singular is a clear cognate. On the (precarious) assumption that these paradigms are cognate paradigms, another connection between the various small paradigmatic structures is attested here. The Manem paradigm does not have any non-singular forms, although the plural marker *kiŋ* can be used to pluralise the reference (Voorhoeve, 1975:416; Foley, 1986:71). A specialised inclusive pronoun is attested in the related languages Amanab (Minch, 1991:31) and Imonda (Seiler, 1985:44), as shown in (6.4b, c)

6.3.3 Up the hierarchy

The paradigmatic structure can be extended up the Horizontal Homophony Hierarchy. The structures that will be considered in this section are shown in Figure 6.4.

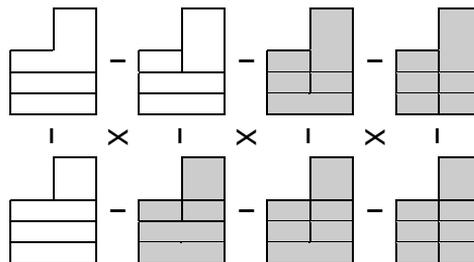


Figure 6.4: The larger paradigms that will be linked to the small paradigms

From the examples that will be presented in this section, it seems to be possible to take big strides through the hierarchy without passing through each stage separately. Once the first-person-hurdle has been taken, the other non-singular forms can follow more easily. Examples from Quechuan, Arawakan, Macro-Gé and Chinese will be discussed in this section. All these families show pronominal paradigms that vary along the Horizontal Homophony Hierarchy, linking the smaller paradigms to the larger paradigms.⁴

Quechuan suffixes

The following case of cognate paradigms comes from the Quechuan languages in Western South America. These paradigms show a connection between the smaller paradigms (that were already discussed previously) and the larger paradigms with more specialised non-singular forms. The paradigms presented all mark the intransitive subject in the various languages.

The paradigmatic structure with the smallest number of oppositions is the paradigm of Tarma Quechua, a Quechua variety spoken in Peru. This paradigm is shown in (6.5a). Plural marking exists, but is not part of the pronominal paradigm (Adelaar, 1977: 89-93, 127-128). A slight departure from this paradigmatic structure is found in Huánuco Quechua. In Huánuco Quechua, the nominal plural marker *...-kuna* is found in the pronominal suffixes with the first person singular only. The resulting paradigm is shown in (6.5b). The first person plural is an integral part of the pronominal paradigm, as it is an aberrant case among the pluralisation strategies of Huánuco Quechua (Weber, 1989: 95-96, 143-144). Both Huánuco and Tarma Quechua are part of subgroup I of the Quechuan languages. The corresponding paradigm in the southern variants of Quechuan (subgroup IIc) is slightly different from these two central Quechua variants. Specialised marking of non-singular is found in all persons in southern Quechua. The suffixes from Bolivian Quechua are shown in (6.5c). This is the most extensive paradigm of the various Quechuan subject paradigms (van de Kerke, 1996: 120-125). All three cases reviewed until now mark an inclusive/exclusive opposition in the paradigm. However, in the northern variants of Quechuan (subgroup IIb), the inclusive form *...-nchi(k)* encompasses the same referential possibilities as the English pronoun 'we'. This is rather different from the other varieties, where the morpheme *...-nchi(k)* has only inclusive reference. The suffixes from Ecuadorian Quechua are shown in (6.5d), as described by Muysken (1977: 43-45).⁵ The suffixes from Inga Quechua, a variant spoken in Columbia (also from subgroup IIb), is slightly different. As can be seen in (6.5e), also an opposition between singular and non-singular in the third person is attested here. The pluralised *...-cuna* is used to mark this opposition; the same morpheme is used in a different function in the paradigm from Huánuco (Schwartz, 1986: 423).

⁴ There are more examples of connections between the various larger paradigms, but these will be discussed in section 6.4 as examples of cognate paradigms that follow the Explicitness Hierarchy.

⁵ Exactly the same paradigm as in Ecuadorian Quechua is also found in Imbabura Quechua (Cole, 1982: 143-145).

(6.5) QUECHUAN

<p>a: Tarma</p> <table border="1" style="margin-left: auto; margin-right: auto;"> <tr><td style="padding: 5px;">...-ñčik</td></tr> <tr><td style="padding: 5px;">...-:</td></tr> <tr><td style="padding: 5px;">...-ñki</td></tr> <tr><td style="padding: 5px;">...-ñ</td></tr> </table>	...-ñčik	...-:	...-ñki	...-ñ	<p>b: Huanuco</p> <table border="1" style="margin-left: auto; margin-right: auto;"> <tr><td style="padding: 5px;">...-nchi:</td></tr> <tr><td style="padding: 5px;">...-:</td><td style="padding: 5px;">...-:kuna</td></tr> <tr><td colspan="2" style="padding: 5px;">...-nki</td></tr> <tr><td colspan="2" style="padding: 5px;">...-n</td></tr> </table>	...-nchi:	...-:	...-:kuna	...-nki		...-n		<p>c: Bolivian</p> <table border="1" style="margin-left: auto; margin-right: auto;"> <tr><td colspan="2" style="padding: 5px;">...-nchik</td></tr> <tr><td style="padding: 5px;">...-ni</td><td style="padding: 5px;">...-yku</td></tr> <tr><td style="padding: 5px;">...-nki</td><td style="padding: 5px;">...-nkichik</td></tr> <tr><td style="padding: 5px;">...-n</td><td style="padding: 5px;">...-nku</td></tr> </table>	...-nchik		...-ni	...-yku	...-nki	...-nkichik	...-n	...-nku
...-ñčik																					
...-:																					
...-ñki																					
...-ñ																					
...-nchi:																					
...-:	...-:kuna																				
...-nki																					
...-n																					
...-nchik																					
...-ni	...-yku																				
...-nki	...-nkichik																				
...-n	...-nku																				
<p>d: Ecuadorian</p> <table border="1" style="margin-left: auto; margin-right: auto;"> <tr><td colspan="2" style="padding: 5px;">...-nchik</td></tr> <tr><td style="padding: 5px;">...-ni</td><td style="padding: 5px;">...-gichik</td></tr> <tr><td style="padding: 5px;">...-ngi</td><td style="padding: 5px;">...-gichik</td></tr> <tr><td colspan="2" style="padding: 5px;">...-n</td></tr> </table>	...-nchik		...-ni	...-gichik	...-ngi	...-gichik	...-n		<p>e: Inga</p> <table border="1" style="margin-left: auto; margin-right: auto;"> <tr><td colspan="2" style="padding: 5px;">...-nchi</td></tr> <tr><td style="padding: 5px;">...-ni</td><td style="padding: 5px;">...-nguishish</td></tr> <tr><td style="padding: 5px;">...-ngui</td><td style="padding: 5px;">...-nguishish</td></tr> <tr><td style="padding: 5px;">...-∅</td><td style="padding: 5px;">...-cuna</td></tr> </table>	...-nchi		...-ni	...-nguishish	...-ngui	...-nguishish	...-∅	...-cuna				
...-nchik																					
...-ni	...-gichik																				
...-ngi	...-gichik																				
...-n																					
...-nchi																					
...-ni	...-nguishish																				
...-ngui	...-nguishish																				
...-∅	...-cuna																				

It will be left to the specialist to propose the precise diachronic development that led to this variable situation (cf Campbell, 1997: 188-189). For the present purpose, it is sufficient to notice that these five paradigmatic structures are closely related and that the similarities are roughly in line with the proposed connections as presented above. The two small paradigms (Tarma Quechua and Huánuco Quechua) are connected to larger paradigms (Bolivian Quechua, and in a different development to Ecuadorian and Inga Quechua). Apparently, these paradigmatic differences can arise in a relatively short time; the morphological changes that led to the presented variation have been quicker than the phonological changes that would lead to large differences in the morphemes themselves. If phonological changes had been more effective, then the correspondences between the morphemes would be much more concealed, making it much harder to track the cognates by the superficial diachronic method that is used here.

Macro-Gé prefixes

The next set of examples of paradigmatic cognates is formed by the pronominal prefixes from various Macro-Gé languages, spoken in Brazil. These languages have variously organised strains of ergativity, which makes it difficult to compare functionally identical paradigms. The following paradigms are chosen because they show unmistakable similarity.

(6.6) MACRO-GÉ

<p>a: Xerente</p> <table border="1" style="margin-left: auto; margin-right: auto;"> <tr><td style="padding: 5px;">wa-...</td></tr> <tr><td style="padding: 5px;">ĩ-...</td></tr> <tr><td style="padding: 5px;">a(i)-...</td></tr> <tr><td style="padding: 5px;">da/∅-...</td></tr> </table>	wa-...	ĩ-...	a(i)-...	da/∅-...	<p>b: Canela-Kraho</p> <table border="1" style="margin-left: auto; margin-right: auto;"> <tr><td style="padding: 5px;">pa-...</td></tr> <tr><td style="padding: 5px;">i-...</td></tr> <tr><td style="padding: 5px;">a-...</td></tr> <tr><td style="padding: 5px;">ih-...</td></tr> </table>	pa-...	i-...	a-...	ih-...	<p>c: Bororo</p> <table border="1" style="margin-left: auto; margin-right: auto;"> <tr><td colspan="2" style="padding: 5px;">pa-...</td></tr> <tr><td style="padding: 5px;">i-...</td><td style="padding: 5px;">xe-...</td></tr> <tr><td style="padding: 5px;">a-...</td><td style="padding: 5px;">ta-...</td></tr> <tr><td style="padding: 5px;">u/∅-</td><td style="padding: 5px;">e-...</td></tr> </table>	pa-...		i-...	xe-...	a-...	ta-...	u/∅-	e-...
wa-...																		
ĩ-...																		
a(i)-...																		
da/∅-...																		
pa-...																		
i-...																		
a-...																		
ih-...																		
pa-...																		
i-...	xe-...																	
a-...	ta-...																	
u/∅-	e-...																	

The first case is the paradigm of the absolutive object prefixes from Xerente, shown in (6.6a). The prefix *wa-...* has ‘unified-we’ reference (Wiesemann, 1986b:365). The other two cases are paradigms of subject prefixes. The subject paradigm from Canela-Kraho is shown in (6.6b). In this case, the morpheme *pa-...* has only-inclusive reference (Popjes & Popjes, 1986:175). The final case is the paradigm from Bororo, as presented in (6.6c). This paradigm has specialised morphemes for almost all non-singular categories (Crowell, 1979:206). The similarity between the Bororo and the Canela-Kraho paradigm is much stronger when compared to the Xerente paradigm. This indicates that the extra morphemes in the Bororo paradigm are probably added relative to the Canela-Kraho paradigm.⁶

Arawakan prefixes

Another group of cognate paradigms is found among the Arawakan languages. The Arawakan languages are spread out over a large part of South America. Still, the subject prefixes as shown in (6.7) present a strong similarity between the widely dispersed languages. Differently from the Macro-Gé case above, the Arawakan examples present a link to a large paradigm without an inclusive/exclusive distinction. Exactly how these differences are to be explained by a diachronic change is a question that must be left to the specialists.

(6.7) ARAWAKAN

a: Ipurina	b: Campa	c: Bare
ni-...	a-...	wa-...
á-...	no-...	nu-...
pi-...	pi-...	bi-... in(i)-...
i-...	ir/o-...	i/wu-... na-...

The first paradigm is the subject paradigm from Ipuriná, an Arawakan language spoken in Brazil. The subject prefixes from Ipuriná are shown in (6.7a). This paradigm has one prefix for all first person plural reference; the prefix *á-...* (Polak, 1894:7). The next case is the subject paradigm from Campa, an Arawakan language spoken in Peru, shown in (6.7b). The individual prefixes are almost identical to the morphemes from Ipuriná, but they form a slightly different paradigmatic structure as the prefix *a-...* has only inclusive reference (Payne, 1981:34; Reed & Payne, 1986:325). Finally, the subject prefixes from Bare, an Arawakan language spoken in Venezuela, are shown in (6.7c). These prefixes are rather different from the other two, although the similarity is strong in the singular forms. The most important difference is the set of overt non-singular forms for the second and third person (Aikhenvald, 1995:27).

⁶ This proposed close link between the paradigmatic structures from Bororo and Canela-Kraho as opposed to Xerente does not coincide with the genetic classification (Campbell, 1997:195-196), nor with the areal distribution (Grimes, 1996:22-23). Impressionistically, this situation seems to be widespread. The comparison of the paradigmatic structure often indicates a slightly different classification from the grouping based on mainly morpho-phonological characteristics.

nections are found. The examples indicate that there is a possible connection between the ‘unified-we’ and the ‘only-inclusive’ paradigms (viz Siouan, Quechuan, Macro-Gé and Arawakan). However, the other diagonal is rather problematic. I did not find any connection between the ‘no-we’ paradigm and the paradigms with an overtly marked inclusive/exclusive opposition (both paradigms are highlighted in Figure 6.5). These two paradigms are too distinct to be connected easily by a diachronic change. Both an inclusive ‘we’ and an exclusive ‘we’ do not seem to be normally added (or lost) at once. The route over one side or the other occurs more easily – adding one morpheme for ‘we’ at a time – and therefore shows up more frequently in a cross-linguistic survey.

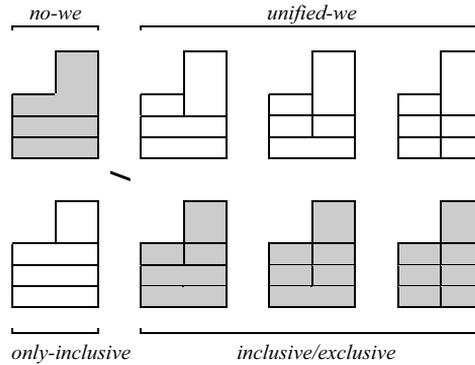


Figure 6.5: Not attested (diagonal) connection between no-we and inclusive/exclusive

A more accurate picture of the possible connections between the various paradigms is obtained when only the structure of the first person complex is taken into account. From this perspective, the finely grained connections from Figure 6.5 are collapsed into more broadly defined associations between paradigmatic structures. The resulting graph is shown in Figure 6.6. The only connection that is not attested is the connection between ‘no-we’ and ‘inclusive/exclusive’.

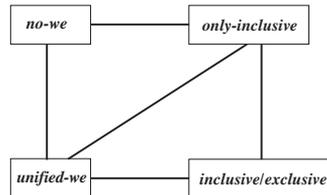


Figure 6.6: Connections attested between various types of the first person complex

6.4 Up and down the explicitness hierarchy

6.4.1 Preamble

The Explicitness Hierarchy fares much better than the Horizontal Homophony when reinterpreted as a hypothesis for diachronic change. The five stages of the Explicit-

ness Hierarchy are repeated in Figure 6.7. As a hypothesis, I propose that pronominal paradigms will develop diachronically along the lines of the hierarchy. If this hypothesis is any good, it is to be expected that there are pronominal paradigms in closely related languages that only differ in one stage on this hierarchy. In this section, I will review the examples that substantiate this hypothesis.

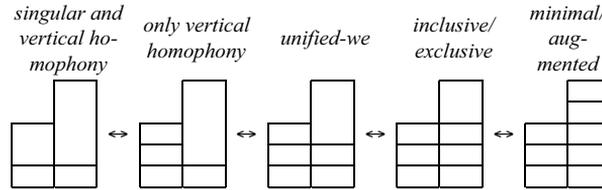


Figure 6.7: The Explicitness Hierarchy

The presentation of the examples will follow the hierarchy from right to left; from more explicit to less explicit. First, in section 6.4.2, the arguments for a connection between the minimal/augmented and the inclusive/exclusive paradigms will be discussed. Second, in section 6.4.3, the link between the inclusive/exclusive and the unified-we paradigms is taken up. Finally, in section 6.4.4, the connections between the unified-we paradigms and the paradigms with singular and/or vertical homophony is discussed.

6.4.2 Minimal/augmented inclusive (or not)

The first step down the explicitness hierarchy is the connection between the minimal/augmented paradigm and the inclusive/exclusive paradigm. This connection is rather strong. This opposition is only added when all other referential categories have already been grammaticalised. Two examples of this connection will be presented, one from the Philippines and one from the Chadic languages. Finally, some areal and structural considerations will be brought forward to show that the minimal/augmented paradigm is not connected to the only-inclusive paradigm.

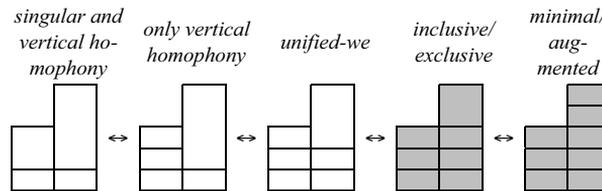


Figure 6.8: Paradigms with or without minimal/augmented opposition

Philippine pronouns

The first case of a set of cognate paradigms on the Explicitness Hierarchy comes from the Philippines. These paradigms present a connection between a completely explicit paradigm (with eight different morphemes) and an almost completely explicit paradigm (with seven different morphemes).

(6.9) PHILIPPINES

a: Cebuano	
	kitá
akú	kamí
ikáw	kamú
siyá	silá

b: Tagalog	
	kata
	tayo
ako	kami
ikaw	kayo
siya	sila

The independent pronouns from Cebuano are shown in (6.9a). There is a full set of non-singular forms, only the two kinds of inclusive reference are not distinguished (Wolff, 1966: 14). The additional opposition between the minimal and the augmented inclusive is found in the closely related language Tagalog, as shown in (6.9b). The pronouns of Cebuano and Tagalog are almost identical. Only the Tagalog morpheme *tayo* is not found in Cebuano.

Chadic pronouns

The Chadic pronouns present another example of cognate paradigms at the high end of the Explicitness Hierarchy. Just as in the Philippine case above, the Chadic pronouns illustrate a connection between the minimal/augmented paradigm (which distinguishes all eight referential categories) and the inclusive/exclusive paradigm (which distinguishes all categories except for the two different inclusions).

(6.10) Chadic

a: Ngizim	
	wà
ná	jà
kà	kwà
∅ (?)	

b: Mandara	
	mà
yà	ɲà
kà	kwà
à	tà

c: Margi	
	-mà
	-mèr
-yù	-'yà
-gù	-nyì
-jà	-ndà

The Ngizim perfective pronouns are shown in (6.10a). The third person plural forms are given as zero in the source, but zero independent third person pronouns are highly suspect (see section 2.6.4). Yet, all other tense/aspect variations of the pronouns also show a homophony between third person singular and plural (Burquest, 1986: 76). The paradigmatic structure of the Mandara completive pronouns, shown in (6.10b), is identical to Ngizim except for the opposition in the third person (Burquest, 1986: 78). Most interesting for the present purpose, Margi adds a minimal/augmented opposition (Hoffmann, 1963: 73-74; Burquest, 1986: 82), as shown in (6.10c).⁸

Disconnected from 'only-inclusive'

These examples from the Philippines and from Chadic show that there is a close connection between the minimal/augmented paradigm and the inclusive/exclusive paradigm. Greenberg (1988: 3-5) proposes another connection: between the minimal/augmented paradigm and the only-inclusive paradigm. Because of certain refer-

⁸ The dashes in (6.10c) indicate that the Margi pronouns are clitics which are added to a root 'nà'.

ential correspondences, this might seem a tempting proposal, but the areal distribution of the paradigmatic structures does not support the idea.

The reasoning by Greenberg goes as follows. The minimal/augmented paradigm (6.11 a) is normally analysed with the inclusive dual 1+2 on a par with the singular persons, as shown in (6.11 b). The second column is now formed by the ‘augmented’ versions of the first column. A natural connection for such a paradigm seems to be a connection to the same paradigms without the augmented forms (6.11 c). Such a paradigm, with only the ‘minimal’ categories, is identical to what has been called ‘only-inclusive’ in the present work, as shown in (6.11 d). Concluding, the minimal/augmented paradigm and the only-inclusive paradigm seem naturally connected.

(6.11)

a:	<table style="border-collapse: collapse; margin: 0 auto;"> <tr><td></td><td style="padding: 2px 10px;">1+2</td></tr> <tr><td></td><td style="padding: 2px 10px;">1+2+3</td></tr> <tr><td style="padding: 2px 10px;">1</td><td style="padding: 2px 10px;">1+3</td></tr> <tr><td style="padding: 2px 10px;">2</td><td style="padding: 2px 10px;">2+3</td></tr> <tr><td style="padding: 2px 10px;">3</td><td style="padding: 2px 10px;">3+3</td></tr> </table>		1+2		1+2+3	1	1+3	2	2+3	3	3+3	=	b:	<table style="border-collapse: collapse; margin: 0 auto;"> <tr><td style="padding: 2px 10px;">1+2</td><td style="padding: 2px 10px;">1+2+3</td></tr> <tr><td style="padding: 2px 10px;">1</td><td style="padding: 2px 10px;">1+3</td></tr> <tr><td style="padding: 2px 10px;">2</td><td style="padding: 2px 10px;">2+3</td></tr> <tr><td style="padding: 2px 10px;">3</td><td style="padding: 2px 10px;">3+3</td></tr> </table>	1+2	1+2+3	1	1+3	2	2+3	3	3+3	~	c:	<table style="border-collapse: collapse; margin: 0 auto;"> <tr><td style="padding: 2px 10px;">1+2, 1+2+3</td></tr> <tr><td style="padding: 2px 10px;">1, 1+3</td></tr> <tr><td style="padding: 2px 10px;">2, 2+3</td></tr> <tr><td style="padding: 2px 10px;">3, 3+3</td></tr> </table>	1+2, 1+2+3	1, 1+3	2, 2+3	3, 3+3	=	d:	<table style="border-collapse: collapse; margin: 0 auto;"> <tr><td></td><td style="padding: 2px 10px;">1+2</td></tr> <tr><td></td><td style="padding: 2px 10px;">1+2+3</td></tr> <tr><td style="padding: 2px 10px;">1</td><td style="padding: 2px 10px;">1+3</td></tr> <tr><td style="padding: 2px 10px;">2</td><td style="padding: 2px 10px;">2+3</td></tr> <tr><td style="padding: 2px 10px;">3</td><td style="padding: 2px 10px;">3+3</td></tr> </table>		1+2		1+2+3	1	1+3	2	2+3	3	3+3
	1+2																																									
	1+2+3																																									
1	1+3																																									
2	2+3																																									
3	3+3																																									
1+2	1+2+3																																									
1	1+3																																									
2	2+3																																									
3	3+3																																									
1+2, 1+2+3																																										
1, 1+3																																										
2, 2+3																																										
3, 3+3																																										
	1+2																																									
	1+2+3																																									
1	1+3																																									
2	2+3																																									
3	3+3																																									

minimal/augmented paradigm in the present work
 minimal/augmented paradigm in the Greenberg analysis
 minimal/augmented paradigm without the augmented forms
 only-inclusive paradigm in the present work

However, the areal distribution of these two paradigms brings a clear end to this strain of thought. The two paradigmatic structures never even come close geographically, nor do they belong to the same genetic family. In the world-wide survey, as described in section 4.7.2, I found four regions where almost all minimal/augmented paradigms are attested. These areas typically include languages with large pronominal systems rather than small ones. The four typical minimal/augmented regions are the Philippines, Nigeria/Cameroon (ie the ‘elbow’ of Africa), California and northwest Australia (‘non-Pama-Nyungan’).⁹ The only-inclusive paradigm is never found in these regions. The only-inclusive type is typically found in completely different areas: New Guinea and South America.¹⁰ This geographical *disconnection* substantiates the typological picture that has been developed above, in which the only-inclusive type is far removed from the minimal/augmented type (cf Figure 6.1). Another argument against a direct connection between the minimal/augmented and the only-inclusive paradigms is found in the morphological structure.¹¹ If there were a set of ‘minimal’ and a set of

⁹ See section 4.7.2 for the complete references; I will here only shortly summarise the attested cases. For the Philippines, see, for example, Tagalog, Maranao and Hanunóo. For the ‘elbow’ of Africa, see, for example, Marghi, Gude, Limbum, Bamileke, Dii and Dan. For California, see, for example, Ute-Southern Paiute, Chemehuevi, Southern Sierra Miwok. Finally, for northwest Australia, see, for example, Marananku, Malakmalak, Bardi, Nyulnyul and Tiwi.

¹⁰ See section 4.5.2 for the complete references; I will here only shortly summarise the attested cases. For New Guinea, see, for example, Nimboran, Imonda and Amanab. For South America, see, for example, Jaqaru, Campa, Maka, Canela-Krahô, Trio and Carib.

¹¹ Greenberg already notes that the morphological structure of the minimal/augmented paradigms does not support his diachronic proposal. He tries to get around this problem by including morphological

‘augmented’ morphemes, one would expect that at least some of these paradigms show indications of a regular morphological derivation of the augmented set. In a sense, the augmented set should look like a regularly derived plural. Of course, not all cases have to show a regular morphological structure, but at least a few should do. However, there is not even one pattern among the various cases of the minimal/augmented paradigm that indicates a regular morphological markedness of the augmented set. This can be illustrated by an example. Blake (1988; 1991) reconstructs a minimal-augmented paradigm for Proto non-Pama-Nyungan, shown here as (6.12). This paradigm turns out to have a regularly derived augmented set by the plural suffix ...-rrV. However, the roots of the minimal and the augmented set are different. The plural marking is not necessary for the distinction. It is probably added as a reinforcement, added only after the complete eight-way paradigm arose.

(6.12) PROTO NON-PAMA-NYUNGAN

	<i>minimal</i>	<i>augmented</i>	
<i>1+2</i>	*nya	*nga-rrV	<i>1+2+3</i>
<i>1</i>	*ngay	*nyi-rrV	<i>1+3</i>
<i>2</i>	*nginy	*nu/ku-rrV	<i>2+3</i>
<i>3</i>	*nu/ngaya	*pu-rrV	<i>3+3</i>

6.4.3 Inclusive/exclusive opposition (or not)

The next step of the Explicitness Hierarchy connects paradigms with an inclusive/exclusive opposition to paradigms without such an opposition (see Figure 6.10).

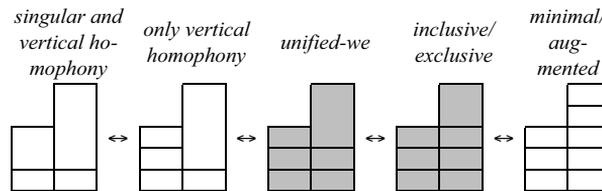


Figure 6.9: the connection between unified-we and inclusive/exclusive

These two kinds of paradigms include all the major variants with some kind of horizontal homophony that were discussed in section 6.3 above. In that section, various cases were presented that linked paradigms with an inclusive/exclusive opposition to paradigms without such an opposition. For the sake of completeness, I add two cases of cognate paradigms here, found among the Western Nilotic pronouns and among the Dravidian pronouns.

Western Nilotic pronouns

An example of this connection can be found in the independent pronouns of the Luo subgroup of Western Nilotic, spoken in Sudan. In (6.13 a), the pronouns from Lango are shown (Bavin, 1981:90). Shown in (6.13 b) are the pronouns from Pāri (Andersen,

[continued from previous page]

elements from outside the pronominal system into the analysis (viz separate number marking). This definitional move is not followed in the present work. A person paradigm is strictly defined as consisting of only prefixes or suffixes.

1988:297). The only paradigmatic difference is the exclusive form in Pāri that is not found in Lango.¹²

(6.13) WESTERN NILOTIC

a: Lango		b: Pāri	
	wan		wání
an		ʔaaní	ʔɔɔní
ym	wun	ʔiiní	ʔúunú
en	gin	yíní	gíní

Dravidian pronouns

Another example is found in the independent pronouns of the Dravidian languages. The Dravidian independent pronouns usually do not have specialised third person forms, using demonstratives instead. The general structure of Dravidian pronouns is exemplified in (6.14b) with pronouns from Malayalam (Asher & Kumari, 1997:226-227,255-258). In Kannada, the exclusive-inclusive distinction that is found in many Dravidian languages, has been lost. The pronouns from Kannada in (6.14a) probably have lost the inclusive-exclusive opposition under influence of the neighboring Indo-European languages (Sridhar, 1990:203).

(6.14) DRAVIDIAN

a: Kannada		b: Malayalam	
	na:vu		nammaɭ
na:nu		ɲaan	ɲaɲɲaɭ
ni:nu	ni:vu	nii	niɲɲaɭ
(demonstr.)		(demonstr.)	

6.4.4 Vertical and singular homophony (or not)

The final step on the Explicitness Hierarchy links the ‘complete’ unified-we paradigm to the paradigms with various kinds of vertical and/or singular homophony. It turns out to be much harder to find examples that illustrate this connection than it has been for the other connections that have been discussed. The reason is probably that there are much fewer paradigms with singular and/or vertical homophony when compared to the abundance of examples of the other paradigmatic structures. Also, the variation is large within the group of paradigms with singular and/or vertical homophony. Still, a few cases of cognate paradigms have been attested, in the Arawakan family, in the Germanic family and possibly in the Gorokan family.

¹² These two languages are only taken to exemplify the pronouns in the Luo languages. Other languages from this family have strongly resembling paradigms. The pronouns from and Anywa (Lusted, 1976:499; Reh, 1996:164) are almost identical to the pronouns from Pāri in. The pronouns from Shiluk (Westermann, 1911:13) and Acholi (Kitching, 1907:9; Crazzolara, 1955:64) are almost identical to the pronouns from Lango.

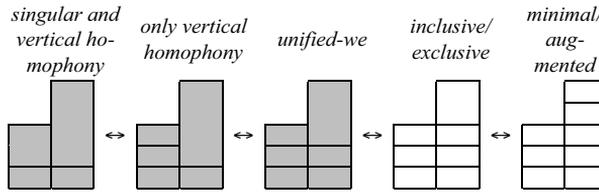


Figure 6.10: Connections between unified-we and different kinds of homophony

Arawakan prefixes

A connection between a unified-we paradigm and a vertical homophony is attested in the Arawakan languages in South America. Some other Arawakan structures were already discussed in section 6.3.3. For the present purpose, the agent prefixes from Bare (Aikhenvald, 1995:27) and Warekena (Aikhenvald, 1998:293) are of interest; they are presented in (6.15). The prefixes from Bare show a complete unified-we paradigm, distinguishing second from third person non-singular. In the cognate paradigm from Warekena, these two forms are identical, possibly a historical merger.

(6.15) ARAWAKAN

a: Bare	b: Warekena												
<table style="border-collapse: collapse; width: 100%;"> <tr> <td style="border: 1px solid black; padding: 5px;">nu-...</td> <td style="border: 1px solid black; padding: 5px;">wa-...</td> </tr> <tr> <td style="border: 1px solid black; padding: 5px;">bi-...</td> <td style="border: 1px solid black; padding: 5px;">in(i)-...</td> </tr> <tr> <td style="border: 1px solid black; padding: 5px;">wu/i-...</td> <td style="border: 1px solid black; padding: 5px;">na-...</td> </tr> </table>	nu-...	wa-...	bi-...	in(i)-...	wu/i-...	na-...	<table style="border-collapse: collapse; width: 100%;"> <tr> <td style="border: 1px solid black; padding: 5px;">nu-...</td> <td style="border: 1px solid black; padding: 5px;">wa-...</td> </tr> <tr> <td style="border: 1px solid black; padding: 5px;">pi-...</td> <td style="border: 1px solid black; padding: 5px;">ni-...</td> </tr> <tr> <td style="border: 1px solid black; padding: 5px;">yu/i/∅-...</td> <td style="border: 1px solid black; padding: 5px;"></td> </tr> </table>	nu-...	wa-...	pi-...	ni-...	yu/i/∅-...	
nu-...	wa-...												
bi-...	in(i)-...												
wu/i-...	na-...												
nu-...	wa-...												
pi-...	ni-...												
yu/i/∅-...													

Germanic suffixes

A well-known case connecting various forms of homophony is found in the Germanic languages. The suffixes shown in (6.16) are the present indicative inflections (dual forms are disregarded) from a few Germanic languages. The Latin suffixes are added in (6.16a) to represent the Proto-Indo-European structure, distinguishing six different morphemes. The Gothic equivalents to the Latin suffixes in (6.16b) still do not show any vertical or singular homophony, although the curious ‘diagonal’ homophony between the third person singular and the second person plural is found here. This ‘diagonal’ homophony is also attested in German (and in Middle Dutch), but in German also a vertical homophony is found. The suffix *...-en* marks both the first and third person plural. Finally, in (6.16d), the Standard Dutch inflection (without inversion) has extended the vertical homophony of the suffix *...-en* to all non-singular referential categories, and has added a singular homophony for addressee and other reference.

(6.16) GERMANIC AND LATIN

a: Latin	b: Gothic	c: German	d: Dutch																						
<table style="border-collapse: collapse; width: 100%;"> <tr> <td style="border: 1px solid black; padding: 5px;">...-o</td> <td style="border: 1px solid black; padding: 5px;">...-mus</td> </tr> <tr> <td style="border: 1px solid black; padding: 5px;">...-s</td> <td style="border: 1px solid black; padding: 5px;">...-tis</td> </tr> <tr> <td style="border: 1px solid black; padding: 5px;">...-t</td> <td style="border: 1px solid black; padding: 5px;">...-unt</td> </tr> </table>	...-o	...-mus	...-s	...-tis	...-t	...-unt	<table style="border-collapse: collapse; width: 100%;"> <tr> <td style="border: 1px solid black; padding: 5px;">...-a</td> <td style="border: 1px solid black; padding: 5px;">...-am</td> </tr> <tr> <td style="border: 1px solid black; padding: 5px;">...-is</td> <td style="border: 1px solid black; padding: 5px;">...-iþ</td> </tr> <tr> <td style="border: 1px solid black; padding: 5px;">...-iþ</td> <td style="border: 1px solid black; padding: 5px;">...-and</td> </tr> </table>	...-a	...-am	...-is	...-iþ	...-iþ	...-and	<table style="border-collapse: collapse; width: 100%;"> <tr> <td style="border: 1px solid black; padding: 5px;">...-e</td> <td style="border: 1px solid black; padding: 5px;">...-en</td> </tr> <tr> <td style="border: 1px solid black; padding: 5px;">...-st</td> <td style="border: 1px solid black; padding: 5px;">...-t</td> </tr> <tr> <td style="border: 1px solid black; padding: 5px;">...-t</td> <td style="border: 1px solid black; padding: 5px;">...-en</td> </tr> </table>	...-e	...-en	...-st	...-t	...-t	...-en	<table style="border-collapse: collapse; width: 100%;"> <tr> <td style="border: 1px solid black; padding: 5px;">...-∅</td> <td style="border: 1px solid black; padding: 5px;">...-en</td> </tr> <tr> <td style="border: 1px solid black; padding: 5px;">...-t</td> <td style="border: 1px solid black; padding: 5px;"></td> </tr> </table>	...-∅	...-en	...-t	
...-o	...-mus																								
...-s	...-tis																								
...-t	...-unt																								
...-a	...-am																								
...-is	...-iþ																								
...-iþ	...-and																								
...-e	...-en																								
...-st	...-t																								
...-t	...-en																								
...-∅	...-en																								
...-t																									

Gorokan pronouns

Another example is found in the Gorokan family, part of the East New Guinea Highlands stock in Papua New Guinea. (Foley, 1986:248-249) reconstructs the independent pronouns from Proto-Gorokan as shown in (6.17). The non-singular shows a vertical homophony. However, none of the Gorokan languages (Foley discusses paradigms from Gende, Siane, Benabena, Kamano-Yagaria and Fore) synchronically displays such a vertical homophony. All languages distinguish second from third person non-singular. So, if this reconstruction is valid, all these Gorokan languages lost the vertical homophony at some time in the past.¹³

(6.17) PROTO-GOROKAN

		*ta
*na		
*ka		*ya
*a		

6.4.5 Summary

The examples presented indicate that the Explicitness Hierarchy fares rather well in the diachronic interpretation. The different stages of the hierarchy are all attested by some examples of cognate paradigms. Examples that jump over different stages of the hierarchy are not attested, at least not in the current set of examples. Of course, different kinds of cognate paradigms might turn up in further research, which would indicate that more possibilities for the diversification of pronominal paradigms exist through time and space. However, from the present data it can be concluded that the Explicitness Hierarchy is a good model to describe some aspect of the diachronic dynamics of pronominal paradigms.

6.5 Conclusion

In this chapter, the hypothesis was tested whether the synchronic restrictions on the structure of pronominal paradigms (as formulated in the previous chapter) can be reinterpreted diachronically. If this turned out to be feasible, then (part of) the explanation of the synchronic restrictions could be attributed to diachronic reasons (cf Plank & Schellinger, 2000). If the structure of human language tends to change in specific directions (for whatever reason), the result of these changes will be shown as a skewed distribution of the possible types in a sample of the world's linguistic diversity.

The method that was used to test the diachronic validity of the synchronic restrictions was to search for 'cognate paradigms'. Cases of such 'cognate paradigms' are pronominal paradigms from genetically closely related languages. Ideally, these pro-

¹³ Foley's argument for this peculiar reconstruction is that the second person plural can only be reconstructed as a compound of first and third plural. When it is taken into consideration that this kind of homophony is rather frequently attested among the non-Austronesian languages in New Guinea, this surprising paradigmatic structure for a reconstruction seems to make sense.

nominal paradigms are functionally identical and the individual morphemes in the paradigm are phonologically closely related. Under these conditions, any differences that might occur in the paradigmatic structure is probably the result of a very shallow change. Examples of such shallow changes present a window on the diachronic connections between the various paradigmatic structures.

Only the major paradigmatic types were included in the present search for cognate paradigms. As was shown in the previous chapter, these major types are structured by two interrelated hierarchies, the Explicitness Hierarchy and the Horizontal Homophony Hierarchy. The Horizontal Homophony Hierarchy did not prove to be a good hypothesis for diachronic change, although some aspects of it could be retained. It turned out that the fine-grained stages of the hierarchy could easily be passed by diachronically. However, a coherent picture emerged when the hierarchy is condensed to the structure of the first person complex only. The Explicitness Hierarchy fares much better as a hypothesis for the diachronic change. From the evidence presented, it can be concluded that this hierarchy is a model for some aspect of the diachronic dynamics of pronominal paradigms. Finally, the remaining connections from the two individual hierarchies can be brought together in one cognitive map of interconnected paradigms, as shown in Figure 6.11. This map is put forward as an informed hypothesis about the lines along which paradigmatic change will happen. More research in detail is needed before the diachronic dimension can be said to be anywhere near being understood. However, the upshot of the present chapter is that a first outline of a theory of paradigmatic change is presented; one that is not based on individual stories of change from incidental examples, but based on the general patterns of pronominal paradigms as attested among the world's languages.

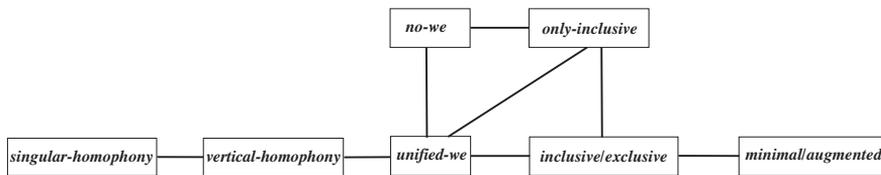


Figure 6.11: Cognitive map of interconnected paradigmatic structures

Part Three

Number Incorporated

“I never ask advice about growing,” Alice said indignantly.

“Too proud?” the other enquired.

Alice felt even more indignant at this suggestion. “I mean,” she said, “that one can’t help growing older.”

“*One* can’t, perhaps,” said Humpty Dumpty; “but *two* can. With proper assistance, you might have left off at seven.”

Lewis Carroll, *Through the looking-glass and what Alice found there*

In this Part Three, I will deal with the subject of number in the domain of person marking. The structure of these chapters is parallel to the discussion in the preceding two parts. First, the cross-linguistic possibilities of number marking are investigated in chapter 7. In this chapter, I will address the variable forms in which pronominal systems mark specific restrictions to the number of participants in a group. Most widely known are special forms for the dual. Other claimed possibilities for marking quantity in human language are the trial (referring to three participants), the quadral (referring to four participants) and the paucal (referring to a few participants). The paucal is a number category that roughly ‘leaves off at seven.’ In chapter 8, a typology is presented of the various paradigmatic structures that show some kind of dual marking. At the end of this chapter, this variation is also analysed quantitatively, which results in the Dual Explicitness Hierarchy. Finally, this hierarchy is interpreted as a guideline for diachronic change. In chapter 9, this diachronic interpretation of the hierarchy is tested by a collection of cognate paradigms. This crypto-diachronic method substantiates the diachronic interpretation of the hierarchy.

Chapter 7

Cardinality

Redefining number in the pronominal domain

7.1 Introduction

The marking of quantity – more traditionally called ‘number’ – in human language is a complex and multifaceted subject. In *The Philosophy of Grammar* Otto Jespersen introduces the discussion of number with the warning that grammatical number is more complex than basic counting.

‘Number might appear to be one of the simplest natural categories, as simple as ‘two and two are four.’ Yet on closer inspection it presents a great many difficulties, both logical and linguistic.’ (Jespersen, 1924:188)

Jespersen continues with an extensive discussion of various aspects of number marking in the grammars of human languages. Among many other things, he notes the problematic status of the plural in the pronominal domain, characterizing the pronoun ‘we’ as being ‘essentially vague’ (Jespersen, 1924:192). Still, he does not conclude that the analysis of ‘we’ as a plural had better be reconsidered. Yet this is exactly what is being proposed in the present work. The traditional use of the concept ‘plural’ in the pronominal domain implies that the number of participants is important. However, as has been argued before (see especially chapter 3), the categories that are traditionally called ‘plural’ are better interpreted as unmarked for number. The difference between the traditional approach and the perspective that is taken in the present work, is illustrated in Figure 7.1.¹

¹ In line with the objective of the present work, I will restrict myself to the marking of number within the pronominal domain, and ignore the marking of number on nouns. There is a powerful universal saying that ‘the differentiation of number with nouns implies that with pronouns’ (Plank, 1989:298). More recent research still substantiates this claim, although more counterexamples have been found (Plank, 1996:124). This claim was first formulated by von Humboldt (1827:156-157), later repeated by Schmidt (1926:316) and by Forchheimer (1953:12). Schmidt puts it thus:

‘Die Numerusbildung ist überall zuerst vom Personalpronomen ausgegangen und hat sich dort am reichsten entwickelt. Es gehört deshalb auch zu den Ausnahmen – die semitischen Sprachen stellen eine solche dar –, daß eine reiche Entwicklung des Numerus beim Nomen vorhanden wäre, aber beim Personalpronomen fehlte.’ (Schmidt, 1926:316)

If this claim is true, then it implies that the present investigation of number within the pronominal domain encompasses also a large part of the variability of number marking on nouns. However, the investigation of the relation between the nominal and the pronominal marking of number falls outside the scope of the present work.

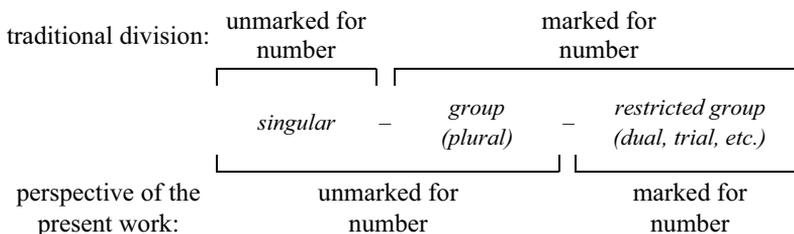


Figure 7.1: Different perspectives on ‘number’ in the pronominal domain

In the previous chapters, the discussion was restricted to pronominal marking that is unmarked for number; viz the marking of singular participants and groups of participants. A group of participants has been defined as a specific combination of various singular participants. The notion ‘group’ is unmarked for number as it is not directed to the number of participants. Of course, a group necessarily refers to more than one person. Yet, the definition of a group is based purely on qualitative characteristics of the constituting parts. The KIND of participants that form the group is important, not the AMOUNT of the included participants. In this chapter, I will turn to cases where amount is marked explicitly in the pronominal domain. The main concept that will be used to describe the marking of amount will be ‘restriction’. Groups, I will argue, can be restricted to the MINIMUM amount of participants, or, eventually, to a SMALL amount. The concepts ‘minimal’ and ‘small’ are proposed as the basic pronominal categories for marking cardinality. In most cases, the concepts ‘minimal amount’ and ‘small amount’ can be freely translated as ‘dual’ and ‘paucal’. All this terminology for the discussion of cardinality may seem to be rather confusing, but the various terms will be used consistently throughout this chapter. The words ‘amount’ and ‘quantity’ are used interchangeably in this chapter; both refer to semantic aspects of cardinality. In contrast, the term ‘number’ will refer to the grammatical form in which aspects of cardinality are grammaticalised in human language. When the phrase ‘amount of X’ is used, this is explicitly intended to mean ‘a specific quantity of X’.²

In this chapter, I will first address the method that will be used to describe the different forms of cardinality in pronominal paradigms. An important aspect of cross-linguistic research is a meta-language on the basis of which the variable forms of human language can be compared. The meta-language for the marking of cardinality will be presented in section 7.2. The complete typology of the various forms of minimally restricted group marking (ie dual marking) that is attested cross-linguistically in pronominal paradigms is too extensive to be dealt with in the context of this chapter. This typology will be postponed to the following chapter. The next section in this chapter deals with the claim that a minimally restricted group (ie dual) is more marked than an unrestricted group (ie plural). Roughly speaking, this is true for pronominal paradigms. Still, there are some counterexamples and problematic cases,

² In many cases, the phrase ‘number of X’ would be a more idiomatic expression for the phrase ‘amount of X’ as it is used here. However, I have not used the word ‘number’ in these cases so as to minimise the confusion with the word ‘number’ as a grammatical term. Incidentally, the phrase ‘number of X’ will be used in the meaning ‘some X’. I have tried to restrict this usage to unambiguous contexts in order not to enlarge this terminological muddle even more.

which will be discussed in section 7.3. The other categories of number – trial, quadral and paucal – turn out to be much less resistant to cross-linguistic comparison than the dual. Their variety is much more restricted compared to the wide variation in marking of the dual. These kinds of marking will be discussed in section 7.4. This chapter will be concluded in section 7.5.

7.2 A meta-language for dual marking

This and the following chapters deal with the marking of number in its various appearances, but most space will be devoted to the marking of the dual. Duals are forms that specifically refer to a group of two participants. They appear in different guises, and are found all over the world. The structural diversity of the dual is large, and equally extensive is the difference between the analyses that are made of the dual in different languages. A complete investigation of the various forms of dual marking in the pronominal domain will be postponed to the next chapter. In this section, the meta-language for the analysis of the dual in pronominal paradigms will be presented. This meta-language is a necessary prerequisite for the comparison of the different patterns of dual marking further on.

I have chosen to use a representation of the dual in pronominal paradigms that is slightly overweighted, but this surplus power will become of use later on. The representation is shown in Figure 7.2. The first two columns consist of the categories as discussed in the previous chapters. The new column is simply a copy of the group categories, labelled as ‘restricted group’.

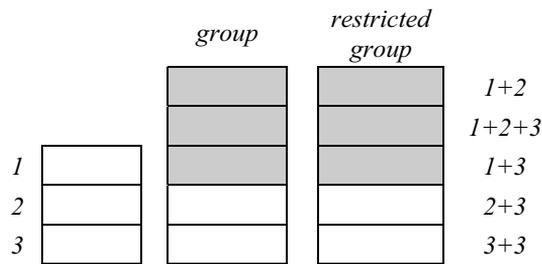


Figure 7.2: Representational scheme for number marking paradigms (the categories to be translated into English as ‘we’ are shaded in grey)

The idea behind the label ‘restricted group’ is that the number of participants in the group is restricted to what is minimally needed.³ For example, to form a group ‘1+3’, at least the speaker and an unspecified set of others is needed. A minimum of two persons is needed: the speaker and only one other participant. In other words, the category ‘restricted 1+3’ is a dual. The restricted forms are mostly dual in nature, but not in all cases. A special case is the category ‘restricted 1+2+3’. This category can never be a dual, as there are at least three persons involved in the category 1+2+3. Traditionally, the category ‘restricted 1+2+3’ is analysed as a trial. Glasgow (1964: 109-110) was

³ The word ‘restricted’ is also used by McGregor in the context of pronominal marking, but in a rather different sense. He uses it to describe the unusual structure of pronominal reference in Gooniyandi, which was discussed in section 3.6.6 (McGregor, 1989:439; 1990: 167-169).

the first to notice that for some languages the name ‘trial’ is referentially correct, while the morphological structure of the language puts this ‘trial’ on a par with the duals. Glasgow used the name ‘dual honored’ for this phenomenon. McKay (1978) later introduced the name ‘unit augmented’. I opted for the name ‘restricted group’ as this label is referentially accurate and it captures both the traditional ‘dual’, and the ‘trial 1+2+3’ form.⁴ The labels ‘group’ and ‘restricted group’ are intended as a proposal to replace the traditional labels ‘plural’ and ‘dual’ in the pronominal domain.

There are two practical issues concerning the representation as shown in Figure 7.2. Contrary to common practice, the ‘restricted group’ column (ie the dual) is placed to the right of the unrestricted ‘group’ column (ie the plural). Normally, the dual is found in between the singular and the plural columns. However, to emphasise the fact that morphemes in the restricted column mark an extra dimension relative to the group forms (viz the quantity of persons in the group), the restricted forms are added sequentially onto the existing framework as it is developed in the previous chapters. This follows the general consensus that the dual is more marked than the plural (see also section 7.3). The second practical issue is that, for reasons of readability, the term ‘dual’ is often used instead of the label ‘restricted group’.⁵ The term ‘dual’ should always be interpreted as meaning ‘minimally restricted group’ in this and subsequent chapters.

As was said before, this framework is slightly overheavy. In the representation in Figure 7.2, there are six different categories that would be translated into English as ‘we’ (shaded grey in the figure). However, these six different forms for ‘we’ are never distinguished morphologically in a paradigm. Five different boxes would suffice to represent the variability of the world’s languages. A better representation for the possible categories would be the structure as shown in Figure 7.3. The two different boxes for 1+2 are joined to one category because the ‘group 1+2’ and the ‘restricted group 1+2’ have the same reference: the pair of speaker and addressee. The representation from Figure 7.3 is referentially sufficient for describing the various forms of the restricted group marking. However, there are numerous pronominal paradigms that emphasise the dual character of their first person pronouns, and in those cases the representation in Figure 7.3 is problematic.

⁴ There is an important difference between the labels ‘restricted group’ and ‘unit augmented’. The label ‘restricted group’ is an etical label (cf footnote 7 on page 228 on the etic/emic distinction). This label is independent of the structure of a particular language. Such language-independent labels are necessary for a comparison of different structures from different languages. In contrast, the label ‘unit augmented’ is an emic label, intended to capture the categorisation that is made by a specific language. The benefits of the label ‘unit augmented’ for the description of the structure of particular pronominal paradigms has been convincingly illustrated for some Australian languages by McKay (1978; 1979; 1984; 1990), but this does not make it effective as the basis for a cross-linguistics comparison (cf section 8.6).

⁵ This improper usage of terminology is not appropriate for the ‘trial 1+2+3’, but this category is grammaticalised only in a small number of cases. The label ‘dual’ will not lead to any misunderstanding until section 8.6, which will deal extensively with this ‘trial 1+2+3’.

	<i>group</i>	<i>restricted group</i>	
			1+2
			1+2+3
1			1+3
2			2+3
3			3+3

Figure 7.3: Alternative representation for the paradigms with restricted group marking

The crux of the problem is the double nature of the category 1+2. This category, traditionally called ‘dual inclusive’, has been aptly named an ‘ambiguous category’ by Greenberg (1988). The label ‘dual inclusive’ refers to a group of two people (dual), including the speaker and the addressee (inclusive). Some pronominal paradigms with a ‘dual inclusive’ mark the linguistic form overtly as dual. In such a paradigm, the ‘dual inclusive’ is surrounded by other dual forms. An example of this marking pattern is found in the independent pronouns from Maori, as shown in (7.1). The dual forms are specifically marked by a suffix *...-ua* (Harlow, 1996:6).

(7.1) MAORI

		<i>group</i>	<i>restricted group</i>	
		tā-tou	tā- ua	1+2
		mā-tou	mā- ua	1+2+3
1	au	kou-tou	kōr- ua	1+3
2	koe	rā-tou	rā- ua	2+3
3	ia			3+3

Other paradigms take a different perspective. The independent pronouns from Umpila, a Pama-Nyungan language from Australia, are structured rather differently, as shown in (7.2). The ‘dual inclusive’ from Umpila does not follow the pattern of the other dual forms that are marked by a suffix *...-ba?amu* (Dixon, 1980:355-356).

(7.2) UMPILA

		<i>group</i>	<i>restricted group</i>	
		ŋali		1+2
		ŋambula		1+2+3
1	ŋayu	ŋana	ŋana- ba?amu	1+3
2	ŋanu	ŋu?ula	ŋu?ula- ba?amu	2+3
3	nhulu	bula	bula- ba?amu	3+3

For languages such as Maori, the ‘dual inclusive’ is explicitly a dual form (that is also inclusive), but for other languages, like Umpila, it is mainly an inclusive form (that is also dual). The referential value is identical in both cases, but the paradigmatic structure that surrounds the ‘dual inclusive’ is different in each case. This ambiguity of the

‘dual inclusive’ can be incorporated by using the representation as shown in Figure 7.2. In that representation, both sides of the ambiguity can be indicated by a different graphic arrangement of the categories. Both variants of the inclusive dual are shown below in Figure 7.4 in the arrangements as they will be represented in the present work. A horizontal combination of categories represents a dual inclusive with emphasis on the inclusivity. A vertical combination of categories represents the dual inclusive, with focus on the duality.⁶

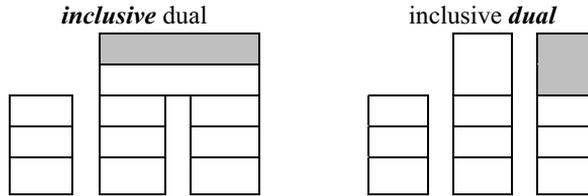


Figure 7.4: The ambiguous nature of the ‘dual inclusive’

A special situation occurs when the emphasis of the ‘dual inclusive’ is on the inclusivity (as shown in the leftmost paradigm in Figure 7.4 above), a. In these cases, the dual inclusive is best analysed along with the singular forms, as shown in Figure 7.5. Of course, referentially, the inclusive dual is never a singular form, but some languages align this dual inclusive structurally with the singulars. In other words, etically it is non-singular, but emically it is a singular.⁷

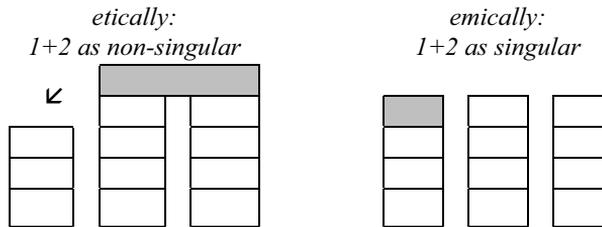


Figure 7.5: The singular-like status of the *inclusive* dual

⁶ In the framework as shown in Figure 7.3, it is not possible to highlight this ambiguous character of the ‘dual inclusive’. Only one of the two natures of the ambiguous ‘inclusive dual’ can be shown in a structure such as Figure 7.3.

⁷ The etic-emic distinction originates from Pike (1954:8-26). Pike’s ETIC approach to language is strongly reminiscent of what in typology has come to be called the ‘tertium comparationis’: the (language-independent) reference structure on the basis of which a cross-linguistic comparison is based. He puts it thus:

‘One of the component goals [of the etic approach] may be said to be nonstructural or classificatory in the sense that the linguist, in using this approach, devises logical categories of systems, classes, and units therein without attempting to make them reflect the structuring discoverable in particular individual languages.’ (Pike, 1954:8)

A typology can be seen as a catalogue and interpretation of the various EMIC instances of the ETIC categories. Each language groups the extra-linguistically defined categories into its own mould.

‘[The emic approach] is an attempt to discover and to describe the pattern of that particular language or culture in reference to the way in which the various elements of that culture are related to each other.’ (Pike, 1954:8)

In the transition as shown in Figure 7.5, the labels of the columns change. In the etical presentation (the leftmost version) the labels are ‘singular’, ‘group’ and ‘restricted group’, respectively, as discussed above. In the emical presentation (the rightmost version), these labels cannot be used because they are referentially incorrect. In the rightmost version, the labels ‘minimal’, ‘augmented’ and ‘unit-augmented’ are more appropriate (McKay, 1978). Only for reasons of cross-linguistics comparison, have I retained the dual inclusive category in its etically justified position in the presentation of the various paradigms that will be shown in the following chapters. However, in all the cases where the inclusive dual is depicted as a horizontal bar, this should be interpreted as signifying that this category is emically singular.

Let me add a last word on the terminology that will be used in the text. As already noted, I will often use the word ‘dual’ where ‘restricted group’ would be more appropriate. In the same sense, the words ‘group’ and ‘plural’ will be used interchangeably in this chapter. Also, the word ‘plural’ is sometimes used in a general sense, encompassing everything that consists of more than one object. Confusingly, this includes the dual. In those instances where I judged the word ‘plural’ to be misleading, I use the term ‘non-singular’ for the general sense of the term ‘plural’.

7.3 Markedness reversals

The analysis of number marking in pronominal paradigms is based on an important assumption. The analysis assumes that restriction to the number of participants in a group is linguistically marked relative to reference to an unrestricted amount. Put more clearly, the analysis starts from the typology of group marking (as established in the previous chapters) and from that point onwards develops an analysis of the marking of number (dual, trial, et cetera). With this method, I suggest that everything there is to say about number marking can be formulated as an addendum to group marking. In this section, I will raise some doubts as to this assumption, only to conclude in the end that it is still safe to proceed this way.

There is a general claim in the literature that the dual is marked relative to general non-singular marking. The basic claim was formulated by Greenberg in one of his well-known universals:

‘Universal 34. No language has a trial number unless it has a dual. No language has a dual unless it has a plural.’ (Greenberg, 1963:94)

Roughly speaking, this is also true within the pronominal domain. Still, there are some counterexamples and problematic cases in which the markedness is reversed (cf Plank, 1989:318). These examples once again highlight the fact that general tendencies among the world’s languages are never to be interpreted too rigorously as universal characteristics. The cases in which the markedness is reversed can be subdivided into three groups. First, there are pronominal paradigms that show REFERENTIAL markedness reversal. Such paradigms have non-singular forms, but the unmarked referential value of these non-singular forms is dual. Second, there are pronominal paradigms that show MORPHOLOGICAL markedness reversal. Such paradigms have both dual and plural forms, but the plural is morphologically derived from the dual. Finally, there are pronominal paradigms that show PARADIGMATIC markedness reversal.

Such paradigms have both dual and plural forms, but there are more categorial distinctions in the dual than in the plural. Each of these cases of markedness reversal will be discussed separately.

The first group of counterexamples to Greenberg's Universal 34 is the set of paradigms that show REFERENTIAL markedness reversal. In these cases, the unmarked non-singular prototypically refers to a group of two persons. Reference to more than two persons can, if necessary, be explicitly marked with an extra morpheme. More specifically, such paradigms have only one non-singular column. These non-singular forms encompass both dual and 'more-than-dual' reference. However, in those contexts where a distinction is needed, it is the reference to 'more-than-dual' that is in need of extra marking. An example of referential markedness reversal is found in the Athabaskan independent pronouns, here exemplified with the pronouns from Navaho (Reichard, 1951:80-89). The form *nxih* can be used meaning 'we' with plural and dual reference. At first glance, this strongly resembles a normal plural, just as the English pronoun 'we' can be used for both plural and dual reference. Interestingly though, in Navaho, when it has to be made explicit that there is a difference between two or more than two persons, then the bare form *nxih* is used for the meaning 'we two'. A prefix *da-...* is used to form the pronoun *da-nxih*, meaning 'we, more than two'. In the description by Reichard, it is put as follows:

'Speakers often fail to distinguish dual and plural, using the same forms for both, unless a distinction is needed, when *da-* is prefixed to the dual forms [to produce plural forms].' (Reichard, 1951:80)

The same situation can be found in the Californian language Achumawi. In the indicative mood, Achumawi uses a suffix *...-má* to disambiguate the plural from the dual. This suffix is not part of the pronominal paradigm, because the pronominal elements are prefixes. The unmarked meaning of the non-singular pronominal prefixes is dual; reference to more than two has to be made explicitly by adding *...-má* (de Angulo & Freeland, 1931:91).⁸

The second group of counterexamples to Greenberg's Universal 34 is the set of pronominal paradigms that show MORPHOLOGICAL markedness reversal. In these cases, the paradigms have different forms for dual and 'more-than-dual' reference. However, when the morphophonological structure of these forms is scrutinised, it turns out that the forms with 'more-than-dual' reference are morphologically marked relative to the dual forms. An example of morphological markedness reversal is found in the independent pronouns from the Uralic language Nganasan. These pronouns are shown in (7.3). The 'more-than-dual' pronouns are derived fully regularly from the dual by a suffix *...-ŋ* (Helimski, 1998:501-502). Note that the morphology reverses the markedness, while the structural markedness is unimpaired: all specialised dual forms have a specialised plural counterpart. There are at least as many group morphemes as there are dual ones.

⁸ See example (8.47) on page 268 for the pronominal prefixes from Achumawi.

(7.3) NGANASAN

		<i>group</i>	<i>restricted group</i>	
		mĩ-ŋ	mi	1+2
1	mənə			1+2+3
2	tənə	tĩ-ŋ	ti	2+3
3	sĩtĩ	sĩtĩ-ŋ	sĩtĩ	3+3

Another example of morphological markedness reversal is found in Damana, an Aruak (Chibchan) language from Colombia. The independent pronouns from Damana are shown in (7.4). The plural pronouns are clearly derived from the dual pronouns by using a suffix *...nyina* (Amaya, 1999:76). For comparison, the independent pronouns of Ika, a language closely related to Damana, are shown in (7.5). The (morphologically unmarked) dual morphemes in Damana are identical to the plural forms in Ika. The forms with the suffix *...nyina* are not attested in Ika (Frank, 1990:26). Probably, Damana has reanalysed the plural pronouns into dual pronouns, and added a new set of plural pronouns with the suffix *...nyina*.

(7.4) DAMANA

		<i>group</i>	<i>restricted group</i>	
		nabi-nyina	nabi	1+2
1	ra			1+2+3
2	ma	mabi-nyina	mabi	2+3
3	na	ijkuna-nyina	ijkuna	3+3

(7.5) IKA

		niwi	1+2
1	nʌn		1+2+3
2	ma	miwi	2+3
3	a	ikʌŋaʔ	3+3

There are some more examples such as the Damana-Ika case, in which the dual forms of a particular language are identical to the plural forms of a close cognate. These examples show that diachronically the dual forms sometimes dominate over the plural ones. An example is the inflectional paradigm from Kwamera, an Oceanic language from Vanuatu as discussed on page 261 below, and the closely related paradigm of Lenakel, as discussed on page 143 above. The dual morpheme *k-...* from Kwamera is found as a general non-singular in Lenakel. The plural from Kwamera is not found in Lenakel. Another example is found in the independent pronouns from two other Oceanic languages, Maori (on page 227 above) and Rapanui (on page 260 below). The

dual morphemes of the second and third person (as still found in Rapanui) have taken over the reference of the plural in Maori.

The third group of counterexamples to Greenberg's Universal 34 contradicts the markedness structure as proposed by the universal. The universal can be interpreted as a markedness cline, for example to be formulated as 'plural is less marked than dual'. In markedness-theory, the occurrence of more oppositions is generally interpreted as a sign of low markedness (Greenberg, 1966:27). Accordingly, paradigms that have categorial oppositions in the dual that are not attested in the plural do not follow this markedness cline. I will call this phenomenon PARADIGMATIC markedness reversal, because the form of the paradigm does not follow the direction of markedness as proposed by the universal. I have found a few examples of paradigmatic counterexamples to Greenberg's Universal 34. All cases have an inclusive-exclusive opposition in the dual but not in the plural. However, in all these cases, the descriptions are not exhaustive and doubt can be raised as to the analysis. Still, as it stands now, these descriptions exist and have to be reckoned with. The first example, shown in (7.6), is the paradigm of the independent pronouns of the Papuan language Samo (Voorhoeve, 1975:391-392).

(7.6) SAMO

		<i>group</i>	<i>restricted group</i>	
		oi	ala	<i>I</i> +2
		nĩ	oli	<i>I</i> +2+3
<i>I</i>	ã	nĩ	oli	<i>I</i> +3
<i>2</i>	nõ	nĩ	nĩli	<i>2</i> +3
<i>3</i>	yõ	yã/diyõ	ili	<i>3</i> +3

Another example with the same structure is the paradigm of the pronominal prefixes from the extinct American language Coos, presented on page 254 below (Frachtenberg, 1922a:321). A final example of paradigmatic markedness reversal is the paradigm of independent pronouns from Wik-Munkan, a Pama-Nyungan language from Australia, presented on page 253. This last example highlights the problematic status of the description of these markedness reversals, as the same author describes slightly different forms for the paradigm on different occasions. One of these descriptions shows paradigmatic markedness reversal; the other does not.

It can be concluded from these examples that reversal of number markedness exists. These examples indicate that not all pronominal paradigms in the world's languages consider the dual to be a marked category relative to the general non-singular. However, the number of examples with markedness reversal is low and it is generally hard to find examples of the kind that are discussed in this section. The cases that were presented in this section highlight that general tendencies in the structure of human language are never simply universal laws. I will disregard these cases in the subsequent chapters and discuss number marking as a specialised variation of group marking.

7.4 Other Numbers

The previous sections have laid the foundation for the discussion of the different ways in which number is marked in pronominal paradigms. The most prolific number-category in the pronominal domain is the dual. However, the variability of dual marking is too extensive to be dealt with here. The next, rather lengthy, chapter is devoted to the description and the analysis of the complete variation of dual marking. I will now turn to other reported categories marking specific amounts. Analogously to the dual, categories such as the ‘trial’ (referring to a group of three), ‘paucal’ (referring to a few) and even an incidental case of a ‘quadral’ (referring to a group of four) have been described in the literature. In this section, I will present a review of the distribution and the structural diversity of the pronominal paradigms with trial, quadral or paucal marking. Although these categories might indeed exist, they are not as widespread as the dual. Also, it is questionable whether the categories ‘trial’ and ‘quadral’ are really strictly used to refer to a specific amount of participants. In many cases (and maybe even in all) ‘paucal’ seems to be a much more useful label.

A major observation on the areal/genetic distribution of these paradigms is that they are found only in the Pacific. I have never come across a pronominal paradigm with a trial, a quadral or a paucal other than among the Austronesian languages or among the non-Austronesian languages from New Guinea and the surrounding islands. Put differently, number marking (other than the dual) seems to be a Melanesian characteristic (Lynch, 1998: 102).⁹ Of course, the linguistic variation in the Melanesian region is immense. Still, this clear geographical restriction gives the impression that these categories are highly marked. The main origin of these categories seems to be the large family of Austronesian languages. These languages are known to have strongly influenced other languages in the Pacific.¹⁰ Only because the trial, quadral and paucal happen to be found in one of the most extended and widespread linguistic families of the world (the Austronesian language family), do these categories appear to be common. Typologically speaking, they are not. Also, the structural diversity of these kinds of number marking is restricted. Most examples of trials, quadrals or paucals simply add a complete set of slightly changed dual forms to the paradigm. Only two incidental cases do not follow this typical structure. This regularity confirms the impression that the marking of higher amounts is a highly constrained phenomenon. The various forms of higher number marking are all morphophonologically related to the dual forms in the paradigm. Consequently, trials, paucals and quadrals will all be interpreted as specialised versions of ‘restricted group’ marking. In section 7.2, a dual has

⁹ I know of one, although poorly described, exception. In the Gé language Apinayé, the pronominal prefix *vamē* is translated as ‘they four’ (Callow, 1962: 115, fn. 4).

¹⁰ Schmidt (1926: 318-320) proposes an influence in the other direction. In his view, extensive number marking originated among the Papuan languages, and was borrowed by the Austronesian languages. This scenario seems less likely once the immense diversity of the Papuan languages is taken into account. Only very few languages out of the many different Papuan languages have a trial or paucal. In contrast, trials are relatively common among the surrounding Austronesian languages. Maybe Schmidt underestimated the diversity of the Papuan languages, wrongly taking the trial of Kiwai, shown here in (7.7), as representative for all Papuan languages. Even so, Schmidt’s section describing the variety of Papuan languages is impressive for his time (1926: 148-154). This indicates that he must have known about the exceptional structure of Kiwai among the Papuan languages.

been analysed as a MINIMALLY restricted group. Extending this proposal, other number categories will be analysed as SPECIFICALLY restricted groups; restricted to either three, four or a few participants. The paradigms of the various number categories will be depicted with extra columns in between the already present dual and group columns; all the columns that mark number are gathered under the heading of ‘restricted group’.¹¹

A paradigm with a trial is attested in the Papuan language Kiwai. The independent pronouns from Kiwai are shown in (7.7) below (Foley, 1986:72, quoting Ray 1933). This is the only example of a complete set of trial forms in a paradigm without an inclusive/exclusive opposition. I have no information on the exact referential possibilities of the trial in Kiwai.

(7.7) KIWAI

		<i>restricted group</i>			
		<i>group</i>	<i>trial</i>	<i>dual</i>	
1	mo	nimo	nimoibi	nimoto	1+2 1+2+3
2	ro	nigo	nigoibi	nigoto	1+3 2+3
3	nou	nei	neibi	neito	3+3

Examples of a trial are numerous among the Austronesian languages. These paradigms always differentiate between an inclusive and an exclusive trial. The dual is diachronically related to the plural plus a suffix meaning ‘two’ and the trial is related to the plural with a suffix meaning ‘three’. This is exemplified in (7.8) with the subject pronouns from the Oceanic language Lonwolwol Ambrym. The dual is formed with the numeral ...-ro, meaning ‘second’ and the third with a suffixed numeral ...-sUl, meaning ‘three’ (Paton, 1971:16,45,47).¹²

(7.8) AMBRYM

		<i>restricted group</i>			
		<i>group</i>	<i>trial</i>	<i>dual</i>	
		er	ensUl	entaro	1+2 1+2+3
1	ni	genem	genemsUl	genemro	1+3
2	nek	gami	gamsUl	gamro	2+3
3	ŋae	ŋe	ŋerUl	ŋero	3+3

¹¹ The division between a minimal inclusive (1+2) and an augmented inclusive (1+2+3) is not attested with higher number marking. Still, I have decided to retain these labels in the diagrams to enhance the visual similarity to the other paradigms that are discussed in the present work.

¹² It is rather easy to find examples of this type among the Oceanic languages. The independent pronouns of Tolai also have a dual and a trial inclusive and exclusive (Mosel, 1984:93). Lynch (1998:102) mentions Anejoñ as another example of the same structure.

It is questionable, though, whether the forms for the dual and the trial are indeed used for precisely two and three persons, respectively. The exact referential possibilities for the dual and trial are taken for granted in most grammars. Only in a few descriptions, is the usage of these categories extensively described. In those cases, it turns out that it is possible to deviate from the exact number of participants. A good example is found in the description of another Oceanic language, Paamese, shown in (7.9) below. The structure of the pronominal paradigm resembles the paradigm from Lonwolwol Ambrym in (7.8). This time, the suffigated numerals are *...-lue* and *...-telu*, meaning ‘two’ and ‘three’ respectively (Crowley, 1982:80-81). Yet, Crowley prefers to call the ‘trial’ by the name ‘paucal’. The following lengthy quote describes extensively how intricately the reference of the trial/paucal is construed.¹³

‘The conditions governing the use of the paucal and the plural are rather more complex. The basic factor that is involved is the absolute size of the group being referred to. Intersecting with this parameter however is the question of relative size, ie whether the group being referred to is contrasted with some larger group within which it is subsumed. When the absolute number is low (say between three and about half a dozen), the paucal is generally used, whether or not there is any contrast with a larger group. (However, the plural will still very occasionally be used even with these low numbers when there is no such contrast.) When the absolute number is in the middle range (say between about half a dozen and a dozen or so), the most significant parameter is that of relative number. For example, one’s own patrilineage will be referred to paucally when it is contrasted with the village as a whole, which will be plural. On the other hand, the patrilineage will be expressed in the plural when contrasted with the nuclear family, which will be in the paucal. As the absolute number increases over the middle range, relative number again becomes less significant, and the plural is generally used for all numbers over a dozen. (However, even with very large numbers, the paucal is occasionally used when the contrast in number is expressed. So while the entire population of Paama will normally be expressed in the plural, even when contrasted with the country as a whole, it has been heard referred to paucally.)’ (Crowley, 1982:80-81)¹⁴

(7.9) PAAMESE

		<i>restricted group</i>			
		<i>group</i>	<i>paucal</i>	<i>dual</i>	
		iire	iatelu	ialue	1+2
		komai	komaitelu	komalu	1+2+3
1	inau	komai	komaitelu	kamalu	1+3
2	kaiko	kamii	kamiitelu	kamilu	2+3
3	kaie	kaile	kaitelu	kailue	3+3

¹³ The independent pronouns of Kwaio are also described using the label ‘paucal’, although the suffix that marks these forms *...-oru* is the numeral three (Keesing, 1985:27-34). Lynch (1998:102) mentions the Nadrau dialect of Fijian as another example of this same structure.

¹⁴ In Paamese, the dual can be used for reference to more than two persons. However, this phenomenon is comparable to the usage of a plural pronoun in addressing a singular person. I will follow Crowley and still use the name ‘dual’ for this category, although this name should not be taken too literally.

‘The dual forms are generally used to refer to only two referents. However, a speaker who is addressing a large crowd, which would normally be addressed in the plural, will sometimes use the dual inclusive forms instead. This is particularly so if the speaker wants to persuade his audience to a particular point of view. The speaker is effectively speaking to each addressee individually, and the use of the dual in such circumstances is more directly appealing to the individual in the large group.’ (Crowley, 1982:80)

Taken at face value, it seems that there are two different types of pronominal systems in the descriptions of Austronesian languages: one with a trial category and one with a paucal (cf Lynch, 1998: 102-103). However, it is unclear whether there really is a difference in linguistic structure between these languages, or whether there is only a difference in descriptive practice. The case for a ‘real’ trial (which would be a morpheme with referential value of exactly three and nothing else) is to my knowledge never explicitly made. When the precise referential possibilities of the trial/paucal are discussed in a grammar, then the vote always goes to paucal reference. Only when no in-depth discussion of the exact referential possibilities is presented, is the name ‘trial’ used. Often, reference will be made to the origin of the trial as a grammaticalised numeral ‘three’. However, in the process of grammaticalisation, the meaning of the morphemes changes frequently. In the case of the numeral ‘three’, the meaning normally changes to ‘few’ in the process of grammaticalisation. This is, for example, argued for the history of the pronouns from the Malaita family of the Solomon Islands. The numeral ‘three’ is grammaticalised into the pronominal paradigm, but in the majority of cases, the meaning changes in this grammaticalisation process from ‘trial’ to ‘paucal’.

‘Only one language, ‘Are’are, has retained the original four number system with a true trial number. Other southern languages, including Kwaio, Sa’a, Langalanga, and Lau, have retained the four number system, but the original trial forms have changed their meaning to “few” or a restricted plural. That the original meaning was “trial” is evident from the fact that the trial pronouns were a compound with the number ‘three’.’ (Simons, 1986:33)

Another case that can shed some light on this issue is the structure of the pronouns from Bolaang Mongondow, spoken on Sulawesi in Indonesia.¹⁵ The pronouns are shown in (7.10) below. The most interesting aspect of these pronouns is a special set of ‘determined’ pronouns that can only be used with an adjacent numeral. This numeral is not restricted to ‘three’; it can in principle, given the right context, be any cardinal number.

(7.10) BOLAANG MONGONDOW

		<i>restricted group</i>		
		<i>group</i>	<i>‘determined’</i>	<i>dual</i>
		kita	kita + numeral	1+2
		kami	kami + numeral	1+2+3
1	aku’oy	kita	kita + numeral	1+3
2	ikow	mo’ikow	kamu + numeral	2+3
3	sia	mosia	tara + numeral	3+3

The dual forms are already highly grammaticalised constructions with the numeral *dua* meaning ‘two’. It is from this situation that a numeral ‘three’ can possibly grammaticalise into the pronominal system. In the change from a numeral to the category of number, the meaning of ‘three’ will probably change to ‘few’.¹⁶

¹⁵ The data and analyses presented from Bolaang Mongondow were kindly given to me in personal communication by Ruben Stoel from Leiden University.

¹⁶ Cf the grammaticalisation of the English *couple* from meaning ‘exactly two’ into meaning ‘a few’.

And what about the quadral? A quadral is said to occur in the pronominal paradigm of the Oceanic language Sursurunga, shown in (7.11). Although Hutchisson (1986:5) uses the label ‘quadral’, his description does not restrict this form to referential usage with groups of four participants only. The ‘quadral’ can also be used for more participants, reminiscent of the use of the paucal as described by Crowell for Paamese. The use of the ‘quadral’ is rather clearly described in the following quote from Hutchisson as being paucal in nature. Corbett (1997) labels this quadral more suitably a ‘greater paucal’.¹⁷

‘Since plural pronouns are never used with relationship terms, the use of these terms skews number reference for both trial and quadral forms (although not for dual) that trial comes to mean a minimum of 3, and quadral a minimum of 4. Either can stand for plural, although quadral is probably more frequently used.’ (Hutchisson, 1986:10)

‘Actual numbers as definitions are less accurate than terms like <few> or <several>, and reflect less the flexibility with which Sursurungas approach their language.’ (Hutchisson, 1986:20, fn.10)

(7.11) SURSURUNGA

		<i>restricted group</i>				
		<i>group</i>	<i>greater paucal</i>	<i>paucal</i>	<i>dual</i>	
		git	gitat	gittul	gitar	<i>1+2</i>
		gim	gimat	gimtul	giur	<i>1+2+3</i>
<i>1</i>	<i>iau</i>	gam	gamat	gamtul	gaur	<i>1+3</i>
		di	diat	ditul	diar	<i>2+3</i>
<i>2</i>	<i>iəu</i>					<i>3+3</i>
<i>3</i>	<i>on,əi</i>					

There are a few paradigms where a trial/paucal is not consistently marked throughout the paradigm. I will only list the two examples attested as an illustration of yet more possible structures for pronominal paradigms. A paucal without an inclusive-exclusive distinction is found in the independent pronouns from the Papuan language Yimas, shown in (7.12). Foley describes the referential properties of the paucal as referring to ‘a few, from three up to seven, but variable depending on context’ (Foley, 1991:111-114). Note that there is a ‘vertical’ homophony in the paucal.

(7.12) YIMAS

		<i>restricted group</i>			
		<i>group</i>	<i>paucal</i>	<i>dual</i>	
		ipa	paŋkt	kapa	<i>1+2</i>
		ipwa	kapwa		<i>1+2+3</i>
<i>1</i>	<i>ama</i>				<i>1+3</i>
<i>2</i>	<i>mi</i>				<i>2+3</i>
<i>3</i>	<i>(locative deixis)</i>				<i>3+3</i>

¹⁷ Schmidt (1926:319) mentions Ambrym and Marina from Vanuatu, as well as Gao from the Solomon Islands, as other examples showing a quadral. I was unable to track down the original sources describing these languages.

Besides the paradigm from Yimas, I have come across one other special paradigm with number marking higher than dual. In the independent pronouns from the Austronesian language Biak, as shown in (7.13), a trial is found only in the third person (Steinhauer, 1985:470,479-480).¹⁸ A comparable situation is found in Bolaang Mongondow, as shown in (7.10) above. In that language, the second and third person plural seem to be more recent innovations.

(7.13) BIAK

		<i>restricted group</i>			
		<i>group</i>	<i>trial</i>	<i>dual</i>	
1	ai'a	ʔo		ʔu	1+2 1+2+3
2	'au	nʔo		nu	1+3
3	i	si/na	sʔo	mu	2+3
				su	3+3

To summarise, number marking higher than dual is neither areally nor structurally very diverse. Areally, the examples are only found in Melanesia. Structurally, almost all patterns are completely regular extensions of one of the standard dual patterns. The only structural exceptions are the patterns from Yimas and Biak as discussed above. More importantly, it remains to be seen whether 'real' trials and quadrals exist. When the description of the referential possibilities is made explicit, it almost always turns out that these categories are better rephrased as 'paucal' and 'greater paucal', even though the historical formation invokes the numeral elements.¹⁹ Historically, these elements clearly mean 'three' or 'four', but once they become grammaticalised into the pronominal paradigm, the meaning changes to paucal and greater paucal.

7.5 Conclusion

In this chapter, the marking of number in the pronominal domain has been redefined around the concept 'restricted group'. Previously (in chapter 3) a GROUP of participants had been defined on a qualitative basis: the kind of participants that form a group is important, not the number. Expanding on this basis, a RESTRICTED GROUP of participants has been defined in this chapter as a group that is marked for the specific number of participants. In this sense, the concept 'restricted group' is intended to replace the traditional concept 'number' in the pronominal domain.

¹⁸ Hein Steinhauer (personal communication) suggest that the pronouns with the ending ...-ʔo are historically related to the numeral three. The referential value of this numeral has been extended to encompass all group reference (cf morphological markedness reversal as discussed in section 7.3). This proposal fits with the observation by Schmidt that many unrestricted plurals in the Austronesian languages are related to former specific number markers:

'Die eine [Bildung] läßt den Plural überhaupt fallen und macht den bisherigen Trial zu einem Plural; das ist die typische Form der sämtlichen polynesischen Sprachen, die aber auch schon in den melanesischen Sprachen auftritt. ... Die andere Bildung fügt zu den Formen Singular, Dual, Trial noch einen Quatral, eine Vierzahl hinzu und macht diese zum Plural.' (Schmidt, 1926:319)

¹⁹ G. Corbett (personal communication) commented that the Austronesian language Larike, spoken on Maluku in Indonesia, has a real grammaticalised trial. I have not been able to check this claim.

Cross-linguistically, the most abundantly attested marking of number is the MINIMALLY RESTRICTED GROUP. In such a group, the number of participants is restricted to the minimally required quantity. For example, the group 1+3 minimally needs two participants (one speaker and one other) for it to retain its characteristic constitution. In most cases, the minimally restricted group is identical to what traditionally has been called a dual. However, this terminology also includes the paradigms that have a minimally restricted 1+2+3 group. This minimally restricted group consists of three participants (one speaker, one addressee and one other) and has been referred to, traditionally, as a trial. The full discussion of the various attested possibilities for the marking of minimally restricted groups in pronominal paradigms will be discussed in the next chapter.

In most cases, the minimally restricted group is marked relative to the unrestricted group. However, there are a few troublesome instances and counterexamples to this markedness cline. Three different kinds of markedness reversal have been discriminated. First, there are cases that show REFERENTIAL markedness reversal. In such cases, the default referential value of an unmarked group morpheme is dual. Second, there are cases that show MORPHOLOGICAL markedness reversal. In these cases, the unrestricted group morphemes are derived morphologically from the restricted forms. Finally, there are cases that show STRUCTURAL markedness reversal. In these cases, the set of restricted group morphemes has more oppositions than the set of unrestricted group morphemes. Notwithstanding these counterexamples, the discussion of the marking of number will be built on the assumption that the restricted group is marked relative to the unrestricted group.

There are various other categories that mark quantity in the pronominal domain. Other categories that have been described in the literature are trials (referring to three participants), quadrals (referring to four participants) and paucals (referring to a few participants). These categories have been attested almost exclusively in the Pacific. The areal and genetic diffusion of these categories is strongly restricted. Moreover, the structural diversity of these paradigms is low. Most cases are simply regular extensions of the paradigms with duals, adding a slightly changed copy of all dual forms to the paradigm. Another point is that, from the existing descriptions, it is questionable whether 'real' trials and quadrals exist. In those case in which a description goes into detail about the referential possibilities of such forms, the conclusion is that 'trials' and 'quadrals' are in fact 'paucals' and 'greater paucals'. The use of the label 'trial' and 'quadral' is only attested in those sources that do not expand on the referential possibilities of these forms. These sources seem to take it for granted that the diachronic source of the affix (often a numeral meaning 'three' and 'four' respectively) is a correct description of the grammaticalised pronominal form. It is doubtful whether this is accurate. Pending more in-depth analyses of the referential possibilities, it seems better to interpret a grammaticalised trial as being in fact a paucal. In the present framework, a paucal can be reformulated as a SMALL restricted group.

Chapter 8

The diversity of minimal number

A typology of dual marking in the pronominal domain

8.1 Introduction

It has often been claimed that duals are most prominent in the marking of ‘we’. The oldest of these observations was made by Humboldt in his paper on the dual:

‘Einigen dieser Sprachen nehmen die Ansicht des Dualis von der redenden und angeredeten Person, dem ICH und dem DU her. In diesen haftet derselbe am Pronomen, geht nur so weit in die übrige Sprache mit über, als sich der Einfluss des Pronomen erstreckt, ja *beschränkt sich bisweilen allein auf das Pronomen der ersten Person in der Mehrheit, auf den Begriff des WIR.*’ (von Humboldt, 1827:156, italics added, MC)

In a thorough reappraisal of Humboldt’s work, Plank (1989) repeats the special status of the first person when it comes to duals. He summarises his findings in two implicational universals:

‘If only one person differentiates a dual, it will very likely be the 1st rather than the 2nd or 3rd. If only two persons differentiate a dual, the 1st is much likelier to be one of them than the 3rd.’ (Plank, 1989:305)¹

The results of this chapter will suggest a slightly different generalisation. Indeed, if there are specialised dual forms in a pronominal paradigm, it is extremely rare for there to be no dual involving at least the speaker. However, when there are grammaticalised dual forms, then it is also extremely rare for there to be no dual involving the other persons as well. Dual forms normally seem to show up across the board of the paradigm.²

Still, there is massive variation in the paradigmatic structure of dual marking. This chapter starts off in section 8.2 with a quick survey of the method that is used to make a typology of this variation and with some short notes on terminology. Subsequently, the various forms of dual marking will be discussed according to the number of dif-

¹ Cf Universal 46 from Sokolovskaya (1980:97).

² There may appear to be a special status for a dual first person, as claimed by Humboldt and Plank, but on closer inspection this apparent predominance is due to the ‘inclusive dual’. When the dual is only attested in the first person inclusive, I have classified such a paradigm as a case of the minimal/augmented paradigm. In languages with a minimal/augmented paradigm, no sign of duality is found in the language except for this ‘inclusive dual’ (cf the data in table 11 from Plank, 1996:131). As argued in sections 3.6.5 and 4.7, the ‘inclusive dual’ is not really a dual, but a special form of group marking. Due to this interpretation, the special status of the first person when it comes to dual marking is strongly diminished (see especially section 8.8.3 below).

ferent forms that include at least the speaker – the number of forms that are to be translated into English as ‘we’. First, I will deal with paradigms with two forms for ‘we’ in section 8.3, leading up to paradigms with five different forms for ‘we’ in section 8.6. Finally, in section 8.7, those paradigms are discussed that have dual forms in the pronominal paradigm, but do not have any dual forms including the speaker. These cases are real counterexamples to the claims by Humboldt and Plank above. In section 8.8, a sample is assembled on the basis of the attested diversity to analyse the structural characteristics of the various paradigmatic possibilities. These quantitative analyses will result in the Dual Explicitness Hierarchy. A summary of the paradigmatic variation attested is presented in section 8.9.

8.2 Method and terminology

In this chapter, the diversity of pronominal paradigms with specialised dual forms will be organised into a typology. For this survey of the various attested structures, the same representation will be used for all examples. This representational structure has been described extensively in the previous chapter; it is repeated here in Figure 8.1. Dual reference in the pronominal domain has been analysed as a ‘minimally restricted group’. A GROUP, as, for example, 1+3, has been defined qualitatively as a set of participants that includes the speaker and others. A MINIMALLY RESTRICTED GROUP 1+3 has been defined as the set of participants that is minimally necessary to form this group – viz two, at least one speaker and one other participant (see chapter 7.2 for a detailed discussion). The minimally restricted groups are in most cases duals, except for 1+2+3, in which case three participants are minimally needed; in other words, the minimally restricted 1+2+3 is a trial. Practically, the labels ‘group’ and ‘restricted group’ are unusual and therefore rather cumbersome. In the running text, the more traditional equivalents ‘plural’ and ‘dual’ will be used regularly.

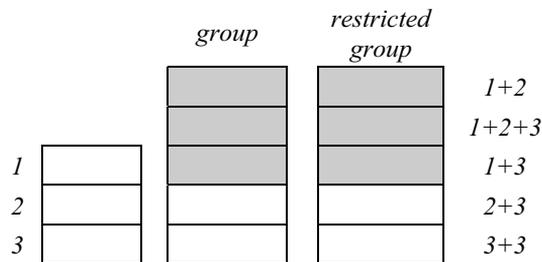


Figure 8.1: Representation for paradigms with dual forms

The discussion of the various paradigmatic structures will be ordered by the number of morphemes that are to be translated into English as ‘we’ (these categories are shaded grey in Figure 8.1). In this way, the resulting typology of the dual can be easily compared with the results obtained in the previous chapters in which the same ‘first person perspective’ was taken. A paradigm that is included in this chapter has minimally two forms for ‘we’ (a plural and a dual form) up to maximally distinguishing five different forms. Some paradigmatic structures will be discussed under the heading of ‘incomplete’ paradigms. The word ‘incomplete’ will consistently be put in quotation-marks to avoid the suggestion that these paradigms are less functional

or deteriorated versions of any ‘complete’ versions. The label ‘incomplete’ is only a result of the classificational strategy that is used for the typology. The classification starts in each section from the paradigmatic structure with maximal distinctiveness and describes other patterns relative to these paradigms. Two characteristics of ‘incomplete’ paradigms will be distinguished: VERTICAL HOMOPHONY – which means that there is an overlap between various categories within one column – and HORIZONTAL HOMOPHONY – which means that there is an overlap between the various columns.

8.3 Two different forms for ‘we’

8.3.1 Preamble

Paradigms with two different forms for ‘we’ come in two guises. One possible division of the two forms for ‘we’ is the inclusive-exclusive opposition. The various paradigmatic structures with such an opposition have already been discussed extensively in chapter 4. The other possible division between the two different forms for ‘we’ is the dual-plural distinction, which will be discussed in this section. These two possible divisions might seem completely different, but they are not. An inclusive ‘we’ prototypically refers to the speech-act dyad of speaker and addressee, which is also a dual. The difference between an inclusive-exclusive opposition and a dual-plural opposition is not that large. A clear criterion to distinguish between the two possible oppositions is needed. The basic criterion for deciding that there is an opposition dual vs. plural (and not inclusion vs. exclusion of the addressee) is the presence of dual forms for all other non-singular categories. The presence of specialised morphemes for the dual second person plural and the dual third person plural indicate that the two forms for ‘we’ are also to be distinguished by number. Such cases, with a full series of dual forms, are discussed in section 8.3.2 under the name of DUAL-UNIFIED-WE paradigm. Next, in section 8.3.3, some examples are presented of paradigms that have a specialised dual ‘we’, but miss some of the other dual categories. There are different kinds of missing referential categories. A difference will be made between HORIZONTAL homophony (different numbers of the same person are taken together in the reference of one morpheme) and VERTICAL homophony (different person of the same number are taken together). Finally, section 8.3.4 will summarise the variation of paradigms with two forms for ‘we’.

8.3.2 The dual-unified-we paradigm

The first paradigmatic structure involving dual marking to be discussed in detail is the dual-unified-we paradigm. Such a paradigm does not distinguish between inclusive and exclusive (viz it has a unified-we structure) and it has dual forms throughout the paradigm. In total, this paradigm consists of nine morphemes: three singular morphemes (1, 2 and 3), three non-singular morphemes (the ‘plural’ forms of the singular categories) and three dual versions of these non-singular morphemes. This is a very common structure for a pronominal paradigm. In his typological analysis of personal pronouns, Ingram (1978:219) classified this paradigm as one of the ‘four systems that are more frequent than the others.’ In this section I will give a review of the worldwide distribution of the dual-unified-we type among the world’s languages. More examples are rather easily found in the areas or families as indicated.

Starting in Europe, a dual pronominal inflection can be reconstructed (partly) for Proto-Indo-European (Szemerényi, 1990:235), but there is hardly a trace of a dual to be found in the pronominal reference of contemporary Indo-European languages. A full dual set is found in Lithuanian (Schmalstieg, 1998:470). In Lower Sorbian, a Slavic language, a complete dual-unified-we paradigm is found in the nominative form of the independent pronouns, shown in (8.1). Note that there is a gender distinction in the third person singular (Stone, 1993:622).

(8.1) LOWER SORBIAN

		<i>group</i>	<i>restricted group</i>	
1	ja	my	mej	<i>I</i> +2
2	ty	wy	wej	<i>I</i> +2+3
3	won(o/a)	woni	wonej	<i>I</i> +3 2+3 3+3

Travelling east through Eurasia, this paradigmatic type is common in Siberian languages of the Uralic family.³ A complete dual-unified-we pronoun system can also be found in Tibeto-Burman.⁴ Crossing the Bering strait to America, duals of this type are attested in Eskimo-Aleut.⁵ However, the pure dual-unified-we pronoun paradigms with duals are not very widespread in the remaining American languages, as most pronominal paradigms with duals also distinguish inclusive from exclusive.⁶ The dual-unified-we pronoun system is also commonly found among the non-Austronesian languages of Papua New Guinea.⁷ Finally, the dual-unified-we paradigm is common in Australia. It is exemplified in (8.2) by the independent pronouns from Warungu (Dixon & Blake, 1979: 11).⁸

(8.2) WARUNGU

		<i>group</i>	<i>restricted group</i>	
1	ŋaya	ŋana	ŋali	<i>I</i> +2
2	yinda	yura	yubala	<i>I</i> +2+3
3	n ^y ula	d ^y ana	bula	<i>I</i> +3 2+3 3+3

³ Eg Khanti (Abondolo, 1998a:373-375) and Mansi (Keresztes, 1998:398-399).

⁴ Eg the independent pronouns from Meithei (Chelliah, 1997:79), Chepang and Kham (Bauman, 1975:273,282).

⁵ Eg the North Slope Iñupiaq suffixes (MacLean, 1986:62-63) and the Central Yup'ic independent pronouns (Woodbury, 1981:231-233).

⁶ Yet, examples of a pure dual-unified-we system can be found. An example is the paradigm of the independent pronouns from Lake/Bodega Miwok (Freeland, 1947:35; Callaghan, 1974).

⁷ Numerous examples can be found in Wurm (1975b) and in Voorhoeve (1975), eg the Bosavi family (ibid. 393-394) and the Duna-Bogaya family (ibid. 396).

⁸ Dixon (1980:329) also mentions Warrgamay and Pitjantjatjara as examples of this paradigmatic type and reconstructs a dual-unified-we paradigms for Proto-Pama-Nyungan (cf Blake, 1988:6).

8.3.3 ‘Incomplete’ dual-unified-we paradigms

Next, I turn to the paradigmatic structure of ‘incomplete’ dual-unified-we paradigms. By using the term ‘incomplete’, I do not want to suggest that these paradigms are less functional or deteriorated variants of the ‘complete’ paradigms. The only reason to approach them under the heading of ‘incompleteness’, is because they will be discussed as variants relative to the dual-unified-we paradigm.

‘Incomplete’ dual-unified-we paradigms come in two guises. First, different persons within the same number can be coded by one and the same morpheme. I will call this VERTICAL homophony, as two categories in the same (vertical) column are combined into the marking of one morpheme. The other kind of homophony, which I will call HORIZONTAL, marks the same person in different numbers by one and the same morpheme. In such paradigms, two categories in the same (horizontal) row are taken together. Ultimately, there are two paradigms with only a dual in the first person and no other duals. These examples are somewhat problematic, as the paradigmatic structure resembles the cases with an inclusive-exclusive distinction. However, the grammatical descriptions of these two examples emphasise the dual character of the distinction, which justifies their inclusion in this chapter. Examples of ‘diagonal’ homophony, in which different persons of different numbers are marked by the same morpheme, are extremely uncommon. The only example that I have come across, the Yareba independent pronouns, is presented at the end of this section.

Vertical homophony is mainly found among the non-Austronesian languages from New Guinea. All these cases from New Guinea have a homophony of the second and third person. As an example of this pattern, the independent pronouns from Kalam are shown in (8.3), with a homophony between second and third in the restricted group but not in the (unrestricted) group marking (Foley, 1986:71, citing Pawley 1966).⁹

(8.3) KALAM

		<i>group</i>	<i>restricted group</i>	
1	yad	cn	ct	<i>1+2</i>
2	nad	nb	nt	<i>1+2+3</i>
3	nwk	ky		<i>1+3</i>
				<i>2+3</i>
				<i>3+3</i>

⁹ The same structure as in Kalam is also found in the Papuan languages Baruya (Wurm, 1975b:499, citing Lloyd 1973) and Fore (Foley, 1986:74, citing Scott 1978). There are different claims about the independent pronouns of Wiru. Wurm (1975b:489, citing Kerr 1967) describes the Wiru paradigm as identical to the structure as found in Kalam in (8.3) but Foley (1986:72, citing Kerr 1966) describes a paradigm identical to the structure as found in Yagaría in (8.4). The same paradigm is also found at the other side of the globe, in the person inflection of the Slavonic languages Slovene (de Bray, 1951:415-416; Priestly, 1993:418) and Upper Sorbian (de Bray, 1951:743; Schuster-Sewc, 1996:161-162).

The homophony between second and third person is often found both in the dual and in the plural (Wurm, 1975b:487; Foley, 1986:71-72). As an example, the past/future suffixes from Yagaria are shown in (8.4) below (Renck, 1975:90-96).¹⁰

(8.4) YAGARIA

		<i>group</i>	<i>restricted group</i>	
1	...-u	...-un	...-uʔ	1+2
2	...-an		...-aʔ	1+2+3
3	...-i	...-a	...-aʔ	1+3
				2+3
				3+3

Interestingly, the opposite situation (homophony in the group marking, but not in the dual) is not attested. This asymmetry substantiates the claim the dual is more marked than the plural. Any opposition that is distinguished in the dual is also found in the plural; or, stated differently, the dual will normally not show a distinction, unless the plural has it (*pace* the cases with ‘paradigmatic’ markedness reversal, see section 7.3). The conflation of second and third person is incidentally extended into the singular. As an example, the ‘past’ suffixes from Siroi are presented in (8.5). The vertical homophony between second and third person is attested here throughout the paradigm (Wells, 1979:30-31).¹¹

(8.5) SIROI

		<i>group</i>	<i>restricted group</i>	
1	...-en	...-geŋ	...-keŋ	1+2
2	...-na	...-naig	...-naik	1+2+3
3	...-na	...-naig	...-naik	1+3
				2+3
				3+3

The only example outside New Guinea of a vertical homophony is found in the Aleut inflection (Geoghegan, 1944:51). This time, the first and third person are coded identically, a pattern not found in New Guinea. The paradigm is shown in (8.6). There are morphophonological different forms of this paradigm, but the homophony between the first and the third person in the dual and the plural is attested in all variants.

¹⁰ The same paradigmatic structure as in Yagaria is also found in Amele, both in the independent pronouns and in the verbal inflection (Roberts, 1987:208), in the inflection of Kewa (Franklin, 1971:57-58) and in the inflection of Kuman (Foley, 1986:70, citing Piau 1985).

¹¹ Exactly the same paradigmatic structure as in Siroi can be found in the suffixes from Magi (Thomson, 1975:631-632).

(8.6) ALEUT

	<i>group</i>	<i>restricted group</i>	
	...-'gin	...-gkin	1+2
1	...-ngan		1+2+3
2	...-mci	...-m'dik	1+3
3	...-'gan	...-gkin	2+3
	...-'gin	...-gkin	3+3

In the case of horizontal homophony, one morpheme is used for different numbers of the same person. An example is the 'preterite' inflection of the Uralic language Mansi, shown in (8.7). Although there are differences between singular, dual and plural in the first and third person, there is only one form for the second person non-singular. This structure is attested in all inflectional person marking in Mansi (Keresztes, 1998:398-407).

(8.7) MANSI

	<i>group</i>	<i>restricted group</i>	
	...-uw	...-əmen	1+2
1	...-əm		1+2+3
2	...-ən	...-en	1+3
3	...-∅	...-iy	2+3
	...-ət		3+3

Another variant of horizontal homophony is described for Gothic, where specialised dual forms only exist for the first and second person. The dual forms are not very often found in the Gothic corpus: 'Sie sind ziemlich selten belegt' (Streitberg, 1920:140). Shown in (8.8) are the suffixes for the 'Präsens Aktiv' of the strong verb *nim*..., meaning 'to take'. This pattern is found throughout the inflection classes of Gothic (Streitberg, 1920:139 ff.).

(8.8) GOTHIC

	<i>group</i>	<i>restricted group</i>	
	...-am	...-ōs	1+2
1	...-a		1+2+3
2	...-iþ	...-ats	1+3
3	...-iþ	...-and	2+3
			3+3

Still another variant of horizontal homophony is found in the independent pronouns of Kewa, a non-Austronesian language from Papua New Guinea, shown below in (8.9). This paradigm has a horizontal homophony between third singular and third dual

(Franklin, 1971:34). The same structure is also found in the Uralic language Nganasan, presented in the previous chapter (see page 231).

(8.9) KEWA

		<i>group</i>	<i>restricted group</i>	
		níáá	sáá	1+2
		nimi	nipi	1+2+3
1	ní			1+3
2	ne			2+3
3	nipú	nimú	nipú	3+3

There are two more cases of horizontal homophony that require some special attention. In these paradigms, there is only one specifically dual form: only for the first person. All other dual reference is taken together with the group marking. As there is only dual marking in the first person complex, these paradigms do not satisfy the basic criterion for dual marking. This criterion dictates that for paradigms to be included in this chapter, there should at least be dual marking outside the first person complex. However, the descriptions of the following two languages below explicitly mention that different forms for ‘we’ are to be distinguished on the basis of the amount of participants. The dual ‘we’ has a preference for an inclusive reading (referring to the speech-act dyad), but this dual ‘we’ can also be shown to have possible exclusive reference in both languages.

(8.10) YIDJŋ

		<i>group</i>	<i>restricted group</i>	
		ŋandŋi	ŋali	1+2
		ŋundu:ba		1+2+3
1	ŋayu			1+3
2	ŋundu			2+3
3	(demonstratives)			3+3

The first of these languages is Yidjŋ, a Pama-Nyungan language from Australia (Dixon, 1977). A specialised dual form is only found in the first person, as shown in (8.10). Dixon explicitly mentions the dual (and not inclusive) meaning of *ŋali*:¹²

‘The main point to note is that *ŋali* forms are seldom used. In fact, *ŋandŋi* is frequently used to refer to two people (one of whom is the speaker). ... *ŋali* is used particularly when it is desired to emphasise TWO people (as opposed to three, etc.). ... [*ŋali*] is most frequent used in an ‘inclusive’ sense, for referring to speaker and addressee ... but it is clear that it can have ‘exclusive’ reference, to speaker and a third party.’ (Dixon, 1977:165-166,179)

¹² The dual form *ŋali* is hardly used. If it would disappear, then the dual will be completely lost in Yidjŋ. The complete loss of the dual has already happened in Tjapukai, a language closely related to Yidjŋ. In the short sketch of Tjapukai by Hale (1976a:237) no mention at all is made of the existence of any dual pronouns.

A similar paradigm is found in the Papuan language Arapesh, shown in (8.11). The paradigm presented is used as the intensive variant of the pronouns, but the other pronominal paradigms in the language have the same paradigmatic structure (Fortune, 1942: 45-49). Just as in Yidj, there is only a dual attested in the first person. There is no dual in the second person, neither is there any dual marking in the class markers or anywhere else in the language. However, in the description it is noted explicitly that the distinction between the two forms for ‘we’ is one of number, not a difference between inclusive and exclusive.

‘There is no distinction between first person exclusive and first person inclusive, such as obtains widely in the languages of the sea peoples of the adjoining Melanesian area.’ (Fortune, 1942:45)

(8.11) ARAPESH

		<i>group</i>	<i>restricted group</i>	
		apak	awhok	$1+2$
		ipak		$1+2+3$
1	eik			$1+3$
2	jak			$2+3$
3	(class markers)			$3+3$

The final example from Yareba combines horizontal and vertical homophony. The independent pronouns from the Papuan language Yareba are shown in (8.12). It looks like a ‘diagonal’ homophony, but it is probably better regarded as a combination of two different kinds of horizontal homophony. Compare it, for example, with a typical Papuan pattern, as the paradigmatic structure of the Kalam pronouns, shown in (8.3) above. Probably, a historical merger resulted in the unique structure as attested in Yareba (Weimer & Weimer, 1975: 675-690).

(8.12) YAREBA

		<i>group</i>	<i>restricted group</i>	
		ya	wa	$1+2$
		ya		$1+2+3$
1	na			$1+3$
2	a			$2+3$
3	dawa	ema		$3+3$

8.3.4 Summary

The ‘complete’ dual-unified-we paradigm is found widespread over the world’s languages. This paradigmatic type is attested in many cases, independently of areal or genetic bonds. Also, a large variety of ‘incomplete’ dual-unified-we paradigm is attested. Paradigms with horizontal homophony as well as paradigms with vertical homophony are found. This situation, characterised by a large amount of variation, is reminiscent of the situation that was attested with the unified-we paradigms (without dual, see section 4.3). In general, paradigms without an inclusive-exclusive opposition are relatively free to show homophony between different parts of the paradigm. Once

a difference between inclusive and exclusive is marked in the paradigm, less variation is found.¹³

Almost all cases with vertical homophony have a homophony between second and third person. Other kinds of vertical homophony are not attested (*pace* the indispensable exception to the cross-linguistic rule). This is rather different from what has been found for paradigms without a dual. Among paradigms without a dual, all possible variants of vertical homophony are attested and the second/third homophony is the least frequent (see section 5.3.3). Also, vertical homophony without duals in the paradigm is attested throughout the world. In contrast, the examples of vertical homophony with a dual are almost all found among the Papuan languages from New Guinea. Finally, for a few of the ‘incomplete’ examples that were presented in section 8.3.3, the homophonous duals are characterised as ‘hardly attested’ or ‘disappearing’ (viz Aleut, Gothic and Yidiñ). At least for these cases, it seems correct to characterise the homophony as a form of deterioration of the paradigm. However, in the other cases, and notably in the Papuan examples, there is no sign indicating that the lack of explicitness in reference is felt as poverty of the system. The term ‘incomplete’ is not to be taken literal in these cases.

8.4 Three different forms for ‘we’

8.4.1 Preamble

The paradigms with three different forms for ‘we’ are a hybrid group. They look alike, but there are slight and possibly important differences between the cases that will be presented in this section. There are two characteristics that limit the various paradigms to be discussed in this section. First, the paradigms have dual forms and, second, they distinguish three different forms for ‘we’. These two characteristics prompted the name ‘dual-3we’ for this type of paradigmatic structure. The dual-3we paradigms which I have found are all incidental exemplars within their close family. Still there are many of these ‘incidents’, indicating that the dual-3we pattern is a clear possibility for the structure of human language.

Before the various examples of dual-3we paradigms will be discussed, a clarification of the defining features is needed. Earlier in the present work, a paradigmatic type was introduced with three different forms for ‘we’: the Maranao-type paradigm with a minimal/augmented opposition in the inclusive (see section 4.7.2). The three forms for ‘we’ in this paradigmatic structure are 1+2 (minimal inclusive), 1+2+3 (augmented inclusive) and 1+3 (exclusive). The paradigms to be discussed in this section are rather different. The main difference is the marking of the dual. The Maranao-type paradigms do not have dual forms for second or third person. In general, the category ‘dual’ is of no importance to these paradigms. In contrast, the dual-3we pronominal paradigms clearly have dual forms. All instances of dual-3we have at least second and

¹³ This observation will be further discussed at the end of section 8.5.3 and in section 8.8.2, after the inclusive-exclusive distinction in the realm of dual marking has been discussed.

third person dual forms. The difference between these two structures is schematically shown in Figure 8.2.

Maranao-type paradigm:

	<i>singular</i>	<i>group</i>	
		3 forms for 'we'	1+2
			1+2+3
1			1+3
2			2+3
3			3+3

Dual-3we paradigm:

	<i>singular</i>	<i>group</i>	<i>restricted group</i>	
		3 forms for 'we'		1+2
1				1+3
2				2+3
3				3+3

Figure 8.2: The difference between minimal/augmented and dual-3we paradigms

In the dual-3we paradigms, the three different forms for 'we' also distinguish dual forms, but it is in this part of the paradigm that the hybrid character of these paradigmatic structures prevails. One possibility is that the dual is only attested in the exclusive (section 8.4.2) or only in the inclusive (section 8.4.3). It is also possible to have an inclusive-exclusive opposition in the dual, but not in the plural (section 8.4.3). Conversely, it is possible to have an inclusive-exclusive opposition in the plural, but not in the dual (section 8.4.5). It is possible that the slight differences between the referential values of the three forms for 'we' are due to different practices in the descriptions, but it is also possible that there are real linguistic differences at stake here. My estimate is that both factors play a role. An argument for the variable descriptions is the case of the Wik-Munkan pronouns, shown in (8.16) below. The same grammarian presents different analyses of this paradigm on different occasions. In contrast, an important argument for the inherent variability of this kind of paradigm will be presented in the next chapter. It will be argued there that the various dual-3we paradigms have cognates showing many different paradigmatic structures. This indicates that the variability of the dual-3we paradigm may be caused by different historical pathways that all lead to this kind of paradigm. If that is the case, then the hybrid character of this paradigmatic type has linguistic saliency. Future in-depth analysis of these languages has to reveal which factors are at stake to explain the attested variability. To preclude a decision on the 'real' structure of these paradigms, I have tried to represent the different descriptions as accurately as possible within the present framework.

8.4.2 Dual in exclusive only

The first example of a dual-3we paradigm, shown in (8.13), is the paradigm of the independent pronouns from Yagua, a genetically isolated language from Peru. The three different forms for 'we' are described as 'exclusive dual', 'exclusive plural' and 'inclusive' (Payne, 1993:20). This is a rather unusual distribution that has not been described before in any typological survey of pronominal systems, as can be inferred from the following claim by Moravcsik (1978):

'No language has been encountered which distinguishes a dual and a plural in the exclusive but not in the inclusive.' (Moravcsik, 1978:352, fn.11)

(8.13) YAGUA

		<i>group</i>	<i>restricted group</i>			
		vúúy			1+2	
		vúúy				1+2+3
1	ráy	núúy	nááy		1+3	
2	jíy	jiryéy	şaşadá		2+3	
3	níí	ríy	naadá		3+3	

Moreover, this unusual paradigmatic structure is not only attested in Yagua. The independent pronouns from the Australian language Ngankikurungkurr are another example, shown in (8.14). In the description of this language, a plural form of the inclusive is added to the paradigm, using a postposition *nime* meaning ‘all’. Strictly speaking, this postposition does not belong to the pronominal paradigm (Hoddinott & Kofod, 1988:94).

(8.14) NGANKIKURUNGKURR

		<i>group</i>	<i>restricted group</i>			
		nayin			1+2	
		nayin				1+2+3
1	ngayi	ngagurr	ngagarri		1+3	
2	nyinyi	nagurr	nagarri		2+3	
3	nem/ngayim	wirrim	wirrike		3+3	

To bring home the point that this pattern is indeed attested, a third example is shown in (8.15), from the Papuan language Savosavo (Todd, 1975:813). Note that both in Savosavo and in Ngankikurungkurr, a gender opposition in the third person singular is present.

(8.15) SAVOSAVO¹⁴

		<i>group</i>	<i>restricted group</i>			
		mai			1+2	
		mai				1+2+3
1	añi	ave	age		1+3	
2	no	me	pe		2+3	
3	lo/ko	ze(po)	to		3+3	

¹⁴ In the short note on Savosavo by Capell (1969a:9-14), the dual forms of the inclusive and exclusive are reversed. This is probably a mistake. Also, a special form *edo mai* is presented for the inclusive dual. However, the word *edo* is the numeral ‘two’ (Todd, 1975:839). Strictly speaking, this word does not belong in the pronominal paradigm.

8.4.3 Dual in inclusive only

The next example of a paradigm with three different forms for ‘we’ comes from the Australian language Wik-Munkan. This example highlights the problematic status of the description of these paradigms, as the same author describes slightly different paradigmatic structures on different occasions. The paradigm shown in (8.16a) comes from Godfrey & Kerr (1964:76) and the paradigm shown in (8.16b) from Godfrey (1964:14). Apart from some slight phonemic changes, and the general drop of the final vowel in (8.16b), the pronominal elements in these paradigms are identical. However, the analysis of the three forms for ‘we’ presents a paradigmatic difference between the two descriptions. The first description in (8.16a), is a case of paradigmatic markedness reversal, as there are two dual forms but only one plural form. In (8.16b), a dual-plural opposition is only found in the inclusive. Whatever the ‘real’ structure of the Wik-Munkan pronouns is, these two paradigms show that it is often difficult to decide for a grammarian what the precise function of the different forms for ‘we’ is.¹⁵

(8.16A) WIK-MUNKAN

		<i>group</i>	<i>restricted group</i>		
		nampi	ngaali	1+2	
			ngana	1+2+3	
1	ngaya		niya	nipa	1+3
2	ninta		tana	pula	2+3
3	nila			3+3	

(8.16B)

		<i>group</i>	<i>restricted group</i>	
		ngamp	ngal	1+2
		ngan		1+2+3
1	ngay	niiy	nip	1+3
2	nint	tan	pul	2+3
3	nil			3+3

8.4.4 Inclusive-exclusive in dual only

The next set of paradigms with three different forms for ‘we’ are just as unusual as the preceding examples. In the following cases, the three forms for ‘we’ are described as ‘inclusive dual’, ‘exclusive dual’ and ‘plural’. The first example, shown in (8.17), is the paradigm of the independent pronouns from the Papuan language Samo (Voorhoeve, 1975:391-392). The second example, shown in (8.18), is the paradigm of

¹⁵ Wik-Munkan is part of the Middle Paman sub-family. In a comparative study of Middle Paman, Hale (1976b:56-57) confirms the analysis from (8.16b). Probably, the analysis as shown in (8.16a) has been a mistake. The description of the Tübatulabal pronouns by Voegelin (1937a:135) present another cases of a comparable mistake (cf footnote 49 on page 156).

the pronominal prefixes from Coos, an extinct language from Oregon (Frachtenberg, 1922a:321).¹⁶

(8.17) SAMO

		<i>group</i>	<i>restricted group</i>	
		oi	ala	1+2
1	ã	nĩ	oli	1+2+3
2	nõ	nĩ	nĩli	1+3
3	yõ	yã/diyõ	ili	2+3

(8.18) COOS

		<i>group</i>	<i>restricted group</i>	
		ɦin-...	is-...	1+2
1	ŋ-...	cɦn-...	xwɦn-...	1+2+3
2	e ^ɛ -...	ɦt-...	ic-...	1+3
3	∅-...	ɦt-...	ux-...	2+3

8.4.5 Inclusive-exclusive in plural only

The next set of examples shows yet another kind of dual-3we paradigm. The descriptions of these languages present, again, a different division of the referential possibilities of the three different forms of ‘we’. This time there is only one dual form, but there is an inclusive-exclusive opposition in the plural. The kind of dual-3we structure appears to be the most widespread variant. The first example, shown in (8.19), is from the Australian language Kuku-Yalanji (Oates & Oates, 1964:7). The second example, shown in (8.20), is the paradigm of the independent pronouns from the Tibeto-Burman language Jiarong (Bauman, 1975:131-132,276). The third example, shown in (8.21), is the paradigm of the independent pronouns from the Papuan language Turaripi, part of the Eleman family on the southeastern tip of Papua New Guinea (Wurm, 1975b:515).

(8.19) KUKU-YALANJI

		<i>group</i>	<i>restricted group</i>	
		ŋana	ŋali(n)	1+2
1	ŋayu	ŋanjın	yubal	1+2+3
2	yuudu	yurra	bula	1+3
3	ñulu	jana		2+3

¹⁶ These paradigms from Samo and Coos were examples were already discussed in section 7.3 as cases of paradigmatic markedness reversal. These paradigms show more forms in the dual than in the plural.

(8.20) JIARONG

		<i>group</i>	<i>restricted group</i>	
		yo	ndžo	1+2
		ŋəñiɛ		1+2+3
1	ŋa	ŋəñiɛ		1+3
2	no	ño	ŋəndʒA	2+3
3	mə	məñiɛ	məndʒAs	3+3

(8.21) TUARIPI

		<i>group</i>	<i>restricted group</i>	
		ereita	elaka	1+2
		ela(o)		1+2+3
1	ara(o)	ela(o)		1+3
2	a(o)	ɛ(o)	euka	2+3
3	are(o)	ere(o)	ereuka	3+3

Two further examples of this pattern come from the Binanderean group. This family of Papuan languages is also found in the southeastern tip of Papua New Guinea, although the Binanderean languages are spoken on the opposite New Guinean coast as the Eleman languages. Shown in (8.22) is the paradigm of the independent pronouns from Guhu-Samane (Richard, 1975: 781). The other example, shown in (8.23), is from Korafe, a language closely related to Guhu-Samane. This paradigm has a typical Papuan vertical homophony between the second and third person. Morphologically, it appears that the exclusive pronoun *namane* is a later addition. The suffix ...-*mane* is consistently glossed as ‘plural’ in the short sketch by Farr (1975: 734-735), although it is never presented as a plural marker on nouns.¹⁷

(8.22) GUHU-SAMANE

		<i>group</i>	<i>restricted group</i>	
		napa	naka	1+2
		nana		1+2+3
1	(a)na	nana		1+3
2	nii	nike	nipe	2+3
3	no	noko	nopo	3+3

¹⁷ The difference between inclusive and exclusive can also be marked in the dual, by preposing either the second person singular pronoun to make *ni nangae* for inclusive reference, or the third person singular pronoun to make *nu nangae* for exclusive reference. These forms ‘rarely occur’ (Farr & Farr, 1975: 734). More importantly, these forms are strictly speaking not part of the pronominal paradigm, but are better seen as some sort of compound of different pronouns.

(8.23) KORAFE

		<i>group</i>	<i>restricted group</i>	
		namonde	nangae	1+2
		namane	nengae	1+2+3
1	na	ne(monde)		1+3
2	ni			2+3
3	nu			3+3

8.4.6 The case of Burarra

Finally, there is one paradigm that also belongs in this hybrid dual-3we type, but it shows yet another distribution of the three forms for ‘we’ over the various referential possibilities. The case in question is the paradigm of the independent pronouns from the Australian language Burarra, as first described by Glasgow (1964). She approaches the description of the pronouns of Burarra in a special way, anticipating what later will be called a ‘unit-augmented’ analysis (see section 8.6). However, two pairs of pronouns in her scheme are identical, leaving only three different forms of ‘we’ (Glasgow, 1964:110-111; 1984:15). Interpreting her description in the framework as used here, the nominative pronouns divide the referential person space as shown in (8.24). The most interesting part is the morpheme *ngatippa*. This pronoun combines the references that are traditionally called ‘inclusive trial’ and ‘exclusive dual’. A highly remarkable combination in traditional terminology, and unusual cross-linguistically (cf the inflection from Bardi and Kunimaipa as discussed in section 3.6.6).¹⁸

(8.24) BURARRA

		<i>group</i>	<i>restricted group</i>	
		ngarripa		1+2
		ngayburrpa	ngatippa	1+2+3
1	ngaypa	anagoyburrpa	anagotippa	1+3
2	nginyipa	birripa	bitipa	2+3
3	nipa			3+3

8.4.7 Summary

The impression that arises from all these examples is one of disorder. It is possible that there are simply many possibilities for the paradigmatic structure, but another option is that some of the descriptions are sloppy with regard to the precise referential properties of the various forms for ‘we’. For most of the examples in this section, the pronominal elements are not extensively documented. Normally, only the paradigm is presented without further arguments for the division of reference. It might well be that the division of the reference between the three forms for ‘we’, once they are studied in

¹⁸ The presented paradigm comes from Glasgow (1984). In this more recent description, Glasgow added feminine variants for the dual forms (not shown here), and made some slight corrections in the spelling of the forms.

more detail, is less variable than appears from these examples. It also might not. A clue that points in the direction of the inherent variability of the dual-3we paradigms is that the examples are almost all incidental cases within their narrow family. Close relatives of the languages that were discussed here show various different paradigmatic types, as will be discussed extensively in the next chapter (see especially section 9.5). This gives the impression that the dual-3we paradigms are an intermediate pattern between other paradigmatic structures. A final observation on the set of dual-3we patterns is that they are mostly found in a restricted area. Nine out of the twelve examples presented are found in southeastern New Guinea and northwestern Australia.

8.5 Four different forms for ‘we’

8.5.1 Preamble

The next pattern to be discussed distinguishes four different forms for ‘we’. In almost all examples to be discussed below, the four forms for ‘we’ arise by a cross-section of the oppositions ‘dual versus plural’ and ‘inclusive versus exclusive’. The most common structure has also a dual-plural opposition for the second and third person. The examples with such a full set of dual forms are discussed in section 8.5.2 under the heading of DUAL-INCLUSIVE/EXCLUSIVE. Next, in section 8.5.3, the paradigms are discussed that do not have a full set of dual forms. The number of ‘incomplete’ patterns with four forms for ‘we’ is much smaller than the ‘incomplete’ patterns with two forms for ‘we’ (cf section 8.3.3). The few examples of ‘incomplete’ paradigms with a dual-inclusive/exclusive paradigm are mostly cases with horizontal homophony: the missing dual reference is taken over by the plural form of the same person. Vertical homophony, as has been attested extensively in dual-unified-we paradigms, is only attested in two exceptional cases here. In section 8.5.4, a special paradigmatic structure with four forms for ‘we’ is discussed. Referentially, these paradigms are identical to the dual-inclusive/exclusive paradigms. However, the morphological structure of these paradigms is rather different; different enough to deserve separate mention and a separate name PARTIAL-UNIT-AUGMENTED. Next, in section 8.5.5, the exceptional case of Palaung is discussed. In Palaung, four forms for ‘we’ are distinguished, but possibly there is no inclusive-exclusive opposition. Instead, a distinction is claimed between ‘present’ and ‘absent’. Finally, section 8.5.6 will summarise the discussion of paradigms with four forms for ‘we’.

8.5.2 The dual-inclusive/exclusive paradigm

The first kind of the various paradigms with four different forms for ‘we’ is the dual-inclusive/exclusive paradigm. This paradigm consists of 11 different morphemes: three singular morphemes (1, 2 and 3), four group morphemes (inclusive, exclusive, second plural and third plural) and dual versions of these four group morphemes. This is a rather common type of paradigmatic structure, characterised by Ingram (1978:219-220) as one of the ‘four systems that are more frequent than the others’. However, outside the Pacific it is not so easy to find examples of this type. In this section, I will present a survey of the distribution of this type of paradigm over the languages of the world.

A complete dual-inclusive/exclusive paradigm is regularly found in Oceanic (Austronesian) languages. Lynch (1998:102) gives examples from Yapese, Nakanai, A'jië and Samoan. It is here exemplified in (8.25) by the pronouns from Maori (Harlow, 1996:6).

(8.25) MAORI

		<i>group</i>	<i>restricted group</i>	
		tātou	tāua	1+2
		mātou	māua	1+2+3
1	au	koutou	kōrua	1+3
2	koe	rātou	rāua	2+3
3	ia			3+3

Such paradigms are also commonly found among the independent pronouns in Australian languages.¹⁹ Outside the Pacific, examples of the dual-inclusive/exclusive paradigm are not so easy to find. There are only a few examples, and they often do not have a complete 11-morpheme system. Either there are gender distinctions in the third person, or the third person morphemes are borderline cases to demonstratives. Some cases are attested among the Tibeto-Burman languages.²⁰ In North America, the dual-inclusive/exclusive paradigm is found rather often.²¹ Finally, an exceptional case of a dual-inclusive/exclusive paradigm is found in Africa. In (8.26) the independent pronouns from Kunama (Bender, 1996:18) are shown. As far as I have seen, this is a unique pattern for an African language. Although it might be an unusual case, it is exceptionally stable. Over a century ago, Reinisch (1881:17) described exactly the same paradigm for the independent pronouns.

(8.26) KUNAMA

		<i>group</i>	<i>restricted group</i>	
		‘kime	‘kiime	1+2
		‘ame	‘aame	1+2+3
1	‘aba	‘eme	‘eeme	1+3
2	‘ena	‘ime	‘iime	2+3
3	‘unu			3+3

¹⁹ Dixon (1980:329,335-336) mentions Dhalandji, Gupapuyŋu, Gumbaynggir, Nyanwaygi and Thargari. Also Warlbiri belongs in this list (Hale, 1973:315-316).

²⁰ In the Tibeto-Burman languages, third person pronouns are often diachronically related to demonstrative elements (Bauman, 1975:107). Yet, Bauman (1975:265-304) lists Bahing, Bunan, Kanauri and Manchati as examples of full dual-inclusive/exclusive paradigms with ‘real’ third person forms.

²¹ On the West Coast, eg Siuslaw (Frachtenberg, 1922b:468) and Chinook (Boas, 1911a:580-581,626). Dual-inclusive/exclusive paradigms are also found among the Iroquoian languages, but with gender marking in the third person, eg the intransitive prefixes from Oneida (Lounsbury, 1953:60-61), Cherokee (Cook, 1979:22) and Tuscarora (Williams, 1976:156). The same structure is found in the pronouns in the Oto-Manguean language Mazahua, which shows gender in the third person (Suárez, 1983:81-82).

8.5.3 ‘Incomplete’ dual-inclusive/exclusive paradigms

There are a few examples of paradigms with four different forms for ‘we’ that are not ‘complete’ dual-inclusive/exclusive paradigms. I will analyse these under the heading of ‘incomplete’ paradigms, although it should be noted that these examples do not show any signs of diachronic instability or deterioration. It is only because I am taking the first person complex as the main classificatory dimension that these examples come out as an ‘incomplete’ subcategory of the dual-inclusive/exclusive paradigm. The ‘missing’ categories in these ‘incomplete’ paradigms will be analysed as cases of either horizontal or vertical homophony. I have predominantly found cases of HORIZONTAL homophony of the dual-inclusive/exclusive paradigm. Such examples combine different numbers of the same person into the marking of one morpheme. The alternative VERTICAL type of homophony – homophony of different person with the same number – is not attested (*pace* the obligatory exception to the cross-linguistic rule).

‘Incomplete’ dual-inclusive/exclusive paradigms are found, for example, in Tibeto-Burman. Shown in (8.27) are the independent pronouns of Limbu. Note the horizontal homophony in the third person (van Driem, 1987:25-28). The same structure (not shown here) is described for the independent pronouns from the closely related language Camling (Ebert, 1997:43).

(8.27) LIMBU

		<i>group</i>	<i>restricted group</i>	
		ani	anchi	1+2
		anige	anchige	1+2+3
1	anɡa	kheni	khenchi	1+3
2	khenɛʔ	khunchi/khenkaʔ		2+3
3	khunɛʔ/khey			3+3

An ‘incomplete’ dual-inclusive/exclusive paradigm with a horizontal homophony in the second person is found in the Australian language Dhuwal. Shown in (8.28) are the nominative pronouns from the Djapu dialect (Morphy, 1983:51-55). The same structure (not shown here) is described for the independent pronouns of Sedang, a Mon-Khmer language from Vietnam (Smith, 1979:80).

(8.28) DHUWAL

		<i>group</i>	<i>restricted group</i>	
		ŋilimurr	ŋali	1+2
		ŋanapurr	ŋiliny(u)	1+2+3
1	ŋarra	nhuma		1+3
2	nhe			2+3
3	ŋayi	walal	manda	3+3

Another example of horizontal homophony in the second person is found in the independent pronouns from Kilivila, an Oceanic language spoken on the Trobriant Islands (Senft, 1986:46-47). Note that the markedness of the four different forms of ‘we’ is

here different from the Limbu case above. In Limbu there seems to be an EXCLUSIVE suffix ...-ge, but in Kilivila there is a PLURAL suffix ...-si (Senft, 1986:31). The Kilivila first person marking is an example of morphological markedness reversal (cf section 7.3).

(8.29) KILIVILA

		<i>group</i>	<i>restricted group</i>	
		yakidasi	yakida	1+2
		yakamesi	yakama	1+2+3
1	yegu	yokwami		1+3
2	yokwa	yokwami		2+3
3	(demonstratives)			3+3

Finally, a horizontal homophony is attested in Rapanui, an Oceanic language from Easter Island (Du Feu, 1996:140). The pronouns are almost identical to the pronouns from Maori in (8.25) above, except for the horizontal homophony in the second and third person. Note that the dual forms for the second and third person have persisted in Rapanui; they have taken over the plural reference. Again, this is a sign that the dual is morphologically sometimes more persistent than the plural (cf the discussion of markedness reversal in section 7.3).

(8.30) RAPANUI

		<i>group</i>	<i>restricted group</i>	
		tatou	taua	1+2
		matou	maua	1+2+3
1	au	korua		1+3
2	koe	korua		2+3
3	ia	raua		3+3

HORIZONTAL homophony is possible in paradigms with a dual, although it is not really widespread among the world's languages. A handful of cases have been shown in this section and another handful of cases (without inclusive-exclusive opposition) has been presented in section 8.3.3. VERTICAL homophony is even less widespread from a cross-linguist perspective. Moreover, the few attested cases of vertical homophony show an interestingly skewed distribution. Vertical homophony is almost exclusively found in paradigms that do not have an inclusive-exclusive opposition. Back in section 8.3.3, many cases of vertical homophony were presented, all without an inclusive-exclusive opposition. Now it turns out that there are no cases of vertical homophony with an inclusive-exclusive opposition (see section 8.8.2 below for a quantitative analysis). This observation can be captured by an implication (8.31).

(8.31) inclusive/exclusive opposition \rightarrow no vertical homophony²²

²² The same implication has been formulated in section 5.3.2. In that section, the implication was formulated in reverse order: 'vertical homophony \rightarrow no inclusive/exclusive opposition'. The reversal as presented here is logically equivalent to that one.

Of course, the world's languages will never let one get away without counterexamples. I have found a vertical homophony in an dual-inclusive/exclusive paradigm in the Papuan language Suena, shown in (8.32). However, it is only found in the inflection of the indicative mood and in the remote tense. Other pronominal paradigms in this language do not show this unusual pattern (Wilson, 1969:97).

(8.32) SUENA

	<i>group</i>	<i>restricted group</i>	
1	...-na	...-nakai	1+2
2	...-sa	...-nakare	1+2+3
3	...-nua	...-wa	1+3
		...-wato	2+3
		...-wato	3+3

Finally, a special paradigmatic structure is found in Kwamera, an Oceanic language from Vanuatu. Kwamera has pronominal subject prefixes on the verb. The subject prefixes are normally followed by a number prefix to disambiguate the reference, but the subject and number paradigms can be separated morphologically (Lindstrom & Lynch, 1994: 12). The subject paradigm is shown in (8.33) below. The exclusive reference is taken care of by the singular form *iak-...* that is used for reference to the speaker. This general use of *iak-...* can be seen as an extended case of HORIZONTAL homophony. In the same paradigm, the dual form *k-...* shows a VERTICAL homophony between the inclusive dual and the third person dual reference. This paradigm from Kwamera is the only example I am aware of that combines a separate horizontal and vertical homophony; a very unusual structure indeed. Some example sentences have been added to exemplify the referential possibilities of these prefixes. In (8.34 a,b), the two different inclusive forms for 'we' are shown. However, (8.34 a) can also refer to a third person dual. The prefix for 'I' and the exclusive 'we' (1+3) are identical, as can be seen in (8.34 c,d).²³

(8.33) KWAMERA

	<i>group</i>	<i>restricted group</i>	
		sa-...	1+2
		k-...	1+2+3
1	iak-...		1+3
2	ik-...		2+3
3	r-...	∅-...	3+3
		k-...	3+3

²³ Compare the Kwamera prefixes with the prefixal paradigm of the closely related language Lenakel, which has been discussed on page 143. In Lenakel, the prefix *k-...* has general plural reference. No dual prefixes are attested in Lenakel.

(8.34) KWAMERA

- a. *k-rou-ánumwi*
 1+2+3/3+3-DUAL-drink
 ‘We two (inclusive) drink’, ‘They two drink’
 (Lindstrom & Lynch, 1994:11)
- b. *sa-ha-iputa*
 1+2+3-PLUR-climb
 ‘We (inclusive) climb’
 (Lindstrom & Lynch, 1994:14)
- c. *ia-pkata-mha*
 1-see-NEG
 ‘I didn’t see it’
 (Lindstrom & Lynch, 1994:14)
- d. *ia-ha-vehe*
 1-PLUR-come
 ‘We (exclusive plural) come’
 (Lindstrom & Lynch, 1994:11)

8.5.4 The partial-unit-augmented paradigm

A different kind of paradigm with four different forms for ‘we’ is a paradigm that I will call PARTIAL-UNIT-AUGMENTED. At first glance, such paradigms are identical to the ‘complete’ dual-inclusive/exclusive paradigms. However, once the morphophonological structure of these paradigms is scrutinised, a slight, but potentially important difference can be observed. As an example of this pattern, the independent pronouns from the Australian language Umpila are shown in (8.35). The dual forms are marked by a special dual suffix *...-ba?amu*. The only dual form that does not use this suffix is the inclusive dual *ηali* (Dixon, 1980: 355-356).

(8.35) UMPILA

		<i>group</i>	<i>restricted group</i>	
		ηambula	ηali	1+2
1	ηayu	ηana	ηana-ba?amu	1+2+3
2	ηanu	ηu?ula	ηu?ula-ba?amu	1+3
3	nhulu	bula	bula-ba?amu	2+3
				3+3

Because of this difference in the morphophonological structure of the dual forms, I propose to represent the Umpila pronouns as shown in (8.36). In this representation, the difference between the two inclusive pronouns *ηali* and *ηambula* is not one of number, but one of person. The opposition DUAL inclusive versus PLURAL inclusive is replaced by an opposition MINIMAL inclusive versus AUGMENTED inclusive. In this way, the connection between these paradigms and other paradigms with a minimal/augmented opposition is emphasised. If the dual forms that are marked by *...-ba?amu* in (8.36) are disregarded, then the paradigmatic structure is of the Maranao-type (see section 4.7.2). In contrast, if the category 1+2+3 also distinguishes a separate restricted group morpheme, then the resulting structure is of the unit-augmented type, which will be discussed in section 8.6. Because of the similarity to

the unit-augmented type, I call the paradigm of Umpila a PARTIAL-UNIT-AUGMENTED paradigm.

(8.36) UMPILA

		<i>group</i>	<i>restricted group</i>	
		ɲali		1+2
		ɲambula		1+2+3
1	ɲayu	ɲana	ɲana-baʔamu	1+3
2	ɲanu	ɲuʔula	ɲuʔula-baʔamu	2+3
3	nhulu	bula	bula-baʔamu	3+3

A somewhat more covert example of this pattern is found in the Australian language Alawa. Shown in (8.37) are the direct pronouns (with a gender distinction in third singular). In this case, there is no overt suffix marking the dual, but a regular alternation between *-l-* and *-ʃ-* (Sharpe, 1972:57).

(8.37) ALAWA

		<i>group</i>	<i>restricted group</i>	
		ñanu		1+2
		ñalu		1+2+3
1	ɲina	ɲalu	ɲaʃu	1+3
2	ñagana	wulu	wuʃu	2+3
3	nuʃa/ɲaduʃa	yiluʃa	yĩʃuʃa	3+3

This kind of paradigms is found in various linguistic families from Australia. Most examples are found among the non-Pama-Nyungan languages in the northwestern part of Australia, except for Umpila that is part of Pama-Nyungan. The partial-unit-augmented paradigm seems to be a typical Australian phenomenon.²⁴

8.5.5 The case of Palaung

All examples with four different forms for ‘we’ that have been discussed until now have the same referential division. The four forms for ‘we’ are always characterised by a cross-section of the oppositions dual-plural and inclusive-exclusive. The only example with a different referential division of the four forms for ‘we’ is found in the Mon-Khmer language Palaung as described by Milne (1921: 17-18). The paradigm of independent pronouns looks like a normal pattern with a dual-plural and an inclusive-exclusive opposition. However, Milne emphasizes that there is no inclusive-exclusive opposition. He describes an opposition between ‘present’ versus ‘absent’ instead. Of course, the present-absent opposition is rather close to the inclusive-exclusive opposi-

²⁴ Examples are attested in Warrwa of the Nyulnyulan family (McGregor, 1994:20-21), in Ngalakan of the Gunwinguan family (Merlan, 1983:71), in Maranungku of the Daly family (Tryon, 1970:16) and in Jaminjung of the Djaminjung family (Schulze-Berndt, 2000:64). Note that both Ngalakan and Maranungku have a gender distinction in the third person singular. A possible case of a partial-unit-augmented paradigms outside Australia is attested in the Gé language Apinayé (see section 9.4.2).

tion. An addressee is always present at the speech act situation, so an inclusive dual is always ‘present’. However, a third person can be either present or absent. Consequently, exclusive marking is crucial to decide between an inclusive-exclusive opposition and a present-absent opposition. On this point, Milne’s description is quite clear. It is present-absent that is marked by the different morphemes, not inclusive-exclusive.²⁵

(8.38) PALAUNG

		<i>group</i>	<i>restricted group</i>	
	ō	ē	āi	<i>we, all present, 1+2 or 1+3</i>
1	mī	yē	yār	<i>we, some not present, 1+3</i>
2	ā̄	pē	pār	<i>you all, 2+3</i>
3		gē	gār	<i>they, 3+3</i>

8.5.6 Summary

The major group of paradigms with four different forms for ‘we’ is the set of DUAL-INCLUSIVE/EXCLUSIVE paradigms. These paradigms have a dual-plural opposition both in the inclusive and in the exclusive; in total there are 11 morphemes in such a paradigm. Examples of the dual-inclusive/exclusive paradigm are commonly found among the world’s languages. Roughly speaking, the main areas/families where this paradigm is found are the Oceanic family, the Pama-Nyungan family, the Tibeto-Burman family and the languages from North America. A special variant of the 11-morpheme paradigm is the PARTIAL-UNIT-AUGMENTED paradigm. The referential structure of this paradigm is identical to the dual-inclusive/exclusive paradigm, but the morphological structure is significantly different. This paradigm is almost exclusively found in Australia. Besides these ‘complete’ patterns, there is a small set of ‘incomplete’ paradigms. Such ‘incomplete’ paradigms are mainly characterised by horizontal homophony; vertical homophony is exceptional. This observation led to the formulation of the implication ‘inclusive/exclusive opposition → no vertical homophony’. Finally, the aberrant case of Palaung was discussed, which looks like a dual-inclusive/exclusive paradigm, but replaces the inclusive-exclusive distinction by a present-absent distinction.

8.6 Five different forms for ‘we’

The last phenomenon that will be tackled under the heading of ‘dual’ marking appears to be a trial, but is better analysed as a twisted kind of dual. It is a pattern with five different forms for ‘we’, exemplified here with the independent pronouns from the Australian language Rembarrnga. The dual-like analysis for Rembarrnga was first

²⁵ In the short description of the Palaung pronouns by Burling (1970:14-17), nothing is found of the present-absent opposition. Instead, he presents a normal inclusive-exclusive opposition. However, the work of Burling is not primarily a grammatical description. Also, he does not refer to the grammar by Milne. I did not find any other sources on Palaung that could shed some more light on the question which analysis is the right one.

proposed by McKay (1978).²⁶ The crucial morpheme is the independent pronoun *ngakorrbbarrah*. This pronoun is traditionally analysed as an inclusive trial, meaning ‘I, you and one other person’. Following the analysis of McKay (1978:28), it seems better to interpret the pronouns of Rembarrnga as an extended version of the Maranao-type paradigm. The Maranao-type paradigm has an opposition between a minimal and an augmented inclusive. The MINIMAL INCLUSIVE is used to refer to the combination of the two principal speech-act participants – speaker and addressee. There is a special pronoun for this category that is paradigmatically put on a par with the singular pronouns. Because this pair is categorised as a singular morpheme, it can be ‘pluralised’. This ‘pluralised’ form is called the AUGMENTED INCLUSIVE. Now, in Rembarrnga, this ‘singular’ pair is not only ‘pluralised’, but also ‘dualised’. The pronoun *ngakorrbbarrah* refers to the speech-act dyad plus one extra person. This group consists of three persons, and therefore, in the referential sense, it is a trial. However, paradigmatically, this form aligns with the other duals, using the suffix *...-bbarrah*. Extending the minimal/augmented terminology, McKay dubbed this phenomenon ‘unit augmented’. His analysis of Rembarrnga is shown in (8.39). Note that there is a gender distinction in the third person singular.

(8.39) REMBARRNGA

	<i>minimal</i>	<i>augmented</i>	<i>unit augmented</i>	
<i>1+2</i>	yukkũ	ngakorũ	ngakorrbbarrah	<i>1+2+3</i>
<i>1</i>	ngũũũ	yarrũ	yarrbbbarrah	<i>1+3</i>
<i>2</i>	kũ	nakorũ	nakorrbbarrah	<i>2+3</i>
<i>3</i>	nawũ/ngadũ	barrũ	barrbbbarrah	<i>3+3</i>

For purposes of comparison and unification, I choose to represent this paradigm slightly differently, as shown in (8.40). I use this graphic representation only to highlight the connection with the other paradigms discussed in this work; I do not intend this representation to have different implications compared to the representation in (8.39). As described in section 7.2, the label ‘unit augmented’ is replaced here with the label RESTRICTED GROUP. The name UNIT-AUGMENTED is used to refer to the paradigmatic structure as a whole.

(8.40) REMBARRNGA

		<i>group</i>	<i>restricted group</i>	
		yukkũ		<i>1+2</i>
		ngakorũ	ngakorrbbarrah	<i>1+2+3</i>
<i>1</i>	ngũũũ	yarrũ	yarrbbbarrah	<i>1+3</i>
<i>2</i>	kũ	nakorũ	nakorrbbarrah	<i>2+3</i>
<i>3</i>	nawũ/ngadũ	barrũ	barrbbbarrah	<i>3+3</i>

²⁶ Glasgow (1964) should be mentioned as a clear precursor of this analysis. Her terminology is a little awkward, but she already had the same ideas as McKay (1978). Compare Glasgow’s analysis of the Burarra independent pronouns, presented on page 256 above.

This paradigmatic pattern is found rather often among the non-Pama-Nyungen languages from northwestern Australia. Many cases of exactly the same paradigmatic structure as shown in (8.40) are attested.²⁷ A special case is the Burarran family. In (8.41) the intransitive prefixes from Djeebbana (McKay, 1978:32) are shown. These prefixes form a complete unit-augmented paradigm. The intransitive prefixes from the related language Burarra are shown in (8.42). These prefixes from Burarra distinguish five different forms for ‘we’, in an unit-augmented-like division (Glasgow, 1984). However, the exclusive non-singular forms are identical to the second person non-singular forms.

(8.41) DJEEBBANA²⁸

		<i>group</i>	<i>restricted group</i>	
		(karr)ka-...		1+2
		ngabarra-...	ngirri-...	1+2+3
1	nga-...	njarra-...	njirri-...	1+3
2	dja-...	narra-...	nirri-...	2+3
3	ka-...	barra-...	birri-...	3+3

(8.42) BURARRA²⁹

		<i>group</i>	<i>restricted group</i>	
		arr-...		1+2
		nguburr-...	arri-...	1+2+3
1	ngu-...	nyiburr-...	nyirri-...	1+3
2	nyi-...	aburr-...	{a}birri-...	2+3
3	{a}			3+3

Outside Australia, this paradigmatic type is almost unattested. There might be unit-augmented systems among the Gé languages in South America, but the examples that I know of are only poorly described.³⁰ The only clear example of a unit-augmented paradigm outside Australia is found in two Papuan languages. The first example is the

²⁷ Rembarnga is part of the Gunwingguan family; other Gunwingguan languages with the same paradigmatic structure of independent pronouns are Ngandi (Heath, 1978:54) and Mangarayi (Merlan, 1982:102). A unit-augmented paradigm is also found in the Gunwingguan language Nungubuyu. However, this paradigm also distinguishes gender in the dual forms (Heath, 1984:241-248). A unit-augmented paradigm is also found in the independent pronouns of the Australian language Nyikina of the Nyulnyulan family (McGregor, 1989:446, citing Stokes 1982).

²⁸ These forms from Djeebbana are the transitive prefixes marking subject with a singular third person object (McKay, 1978:32). I expect these to be rather close to the intransitive forms. However, McKay (1984:144) remarks in a later work that these ‘earlier tabulations need some correction and completion ... A more complete discussion is in preparation.’

²⁹ This vertical homophony in the inflectional paradigm of Burarra can be disambiguated by adding an independent pronoun. However, the independent pronouns are of another atypical paradigmatic structure with homophony between 1+2+3 and 1+2, presented on page 256 above. The independent pronouns as well as the pronominal inflection allow some representational ambiguity, but together they function as a full unit-augmented system.

³⁰ Maybe Kayapo has a unit-augmented paradigm (Wiesemann, 1986b:368-369). See section 9.4.2 for a discussion of this pronominal marking in the Gé languages.

Papuan language Weri. The independent pronouns from Weri are shown in (8.43). In this case, the dual forms are regularly derived from the group morphemes, by using a suffix *...-ip* (Boxwell, 1967:36). The language Kunimaipa is a close relative of Weri; both are part of the Goilalan family. The independent pronouns from Kunimaipa are shown in (8.44). The paradigmatic structure turns out to be rather unusual. Strictly speaking, there are only four different forms for ‘we’ in Kunimaipa (*rei*, *rei-pi*, *rari* and *rari-pi*). However, the referential values of these four forms can only be understood if the paradigm is analysed as a variant of the unit-augmented paradigm. The form *rei-pi* is a general dual form, used for both inclusive and exclusive dual. The form *rari-pi* is the odd one out: it used only for restricted 1+2+3 (Pence, 1968; Geary, 1977: 17-18).

(8.43) WERI

		<i>group</i>	<i>restricted group</i>	
		tepir		1+2
		tëar	tëar-ip	1+2+3
1	ne	ten	ten-ip	1+3
2	në	ar	ar-ip	2+3
3	pä	pët,pëar	pëar-ip	3+3

(8.44) KUNIMAIPA³¹

		<i>group</i>	<i>restricted group</i>	
		rei-pi		1+2
		rari	rari-pi	1+2+3
1	ne	rei	rei-pi	1+3
2	ni	ari	ari-pi	2+3
3	pi	paru	paru-pi	3+3

The other example outside Australia is described for the East Papuan language Reefs, spoken in the Solomon Islands. Again, there is a clear suffix for the ‘restricted group’ forms; this time the suffix is *...-le* (Wurm, 1969:83).

(8.45) REEFS

		<i>group</i>	<i>restricted group</i>	
		...-d ^y i		1+2
		...-de	...-dele	1+2+3
1	...-nɔ	...-ŋo	...-ŋole	1+3
2	...-mu	...-mi	...-mile	2+3
3	...-gu	...-gui	...-guile	3+3

³¹ The use of the suffix *...-pi* in Kunimaipa is said to be ‘optional’ (Geary, 1977:17). It is added here to show the resemblance with the Weri paradigm. Without the suffix *...-pi*, the structure of the independent pronouns in Kunimaipa would be of the same type as the exceptional structure that is also found in the independent pronouns from Gooniyandi and Yaouré, as described in section 3.6.6.

To conclude, the unit-augmented paradigm is a typical Australian phenomenon. There are only a few cases outside Australia, and those are found mainly closely surrounding Australia. The same geographical bias is also found in the partial-unit-augmented patterns, discussed previously in section 8.5.4. As the name indicates, not only the geographical distribution, but also the paradigmatic structure of the partial-unit-augmented paradigm is reminiscent of the unit-augmented paradigms discussed here. The relation between these two patterns is further investigated in the next chapter (see especially section 9.4).

8.7 Other duals

There are three different kinds of paradigms with duals outside the first person complex. The dual can be found only in the second person; the dual can be found only in the third person; or finally, the dual can be found both in the second and third person. There are examples of all three cases described in the literature, although they are not numerous, nor all equally convincing. Still, this pattern is a clear possibility for the structure of human language.

I know of one example of a paradigm with only a second person dual. The subject pronouns from the Omotic language Dizi, shown in (8.46), include a second person dual, besides a second person plural pronoun (Allen, 1976:383). Earlier descriptions do not mention the dual (Allen, 1976:392, fn.6). Note that the third person singular shows a gender distinction.

(8.46) DIZI

		<i>group</i>	<i>restricted group</i>		
1	yinu	inu			1+2 1+2+3
2	yetu	iti		it	1+3 2+3
3	iti / iži	iži			3+3

I know of one clear example of a sole third person dual, from the American language Achumawi. The verb prefixes in the indicative mood from Achumawi are shown in (8.47).

(8.47) ACHUMAWI

		<i>group</i>	<i>restricted group</i>		
		h-...			1+2 1+2+3
1	s-...	s.h-...			1+3
2	k-...	gidz-...			2+3
3	y-...	y-...		eiy-...	3+3

In Achumawi, a verb suffix *...-má* is used to disambiguate plural from dual, but for the third person there is a separate dual prefix, and the suffix *...-má* is not used (de Angulo & Freeland, 1931:91).³² There are a few more languages that have pronominal paradigms with only duals in the third person, but all these have gender distinctions as well. I just mention them shortly for completeness. Another American language, Tunica, has a separate third person dual in the independent pronouns and in the prefixal pronominal paradigm (Haas, 1946:347-363). In Europe, the independent pronouns of Slovene have a third person dual (Priestly, 1993:408-409). Finally, on the opposite side of the world, the same phenomenon is described for the independent pronouns of the Papuan language Buin (Wurm, 1975a:794, citing Laycock 1978).³³

Duals in the second and third person are found in Aleut. Shown in (8.48) are the independent pronoun in the nominative (Geoghegan, 1944:32).³⁴

(8.48) ALEUT

		<i>group</i>	<i>restricted group</i>	
		tuman		1+2
		tuman		1+2+3
1	ting			1+3
2	txin	txici	txi'dik	2+3
3	ingan	ingakun	ingaku	3+3

Finally, mention should be made in this context of the inflectional suffixes from the Papuan language Kapau. Kapau has a few different pronominal paradigms in the indicative mood. The suffixes for the Immediate Future distinguish 7 different forms, which form a typical Papuan paradigm such as the one from Yagaria in (8.4) above. Other tenses have reduced this paradigm to rather strange patterns, involving dual in non-first person only (Oates & Oates, 1969:26-35). To conclude, a dual outside the first person complex is indeed a real possibility for a pronominal paradigm, although it does not seem to be a pervasive phenomenon.

³² The suffix *...-má* is used to mark the plural. The unmarked form of the non-singular prefixes have dual reference. This is a case of referential markedness reversal (see section 7.3).

³³ Plank (1989:304) also mentions southern Arabic varieties as examples of languages that only have a dual in the third person.

³⁴ Plank (1989:303) mentions ancient Greek (which lost the dual, starting in the first person) and Classic Arabic as examples with only a second and third person dual. Both language have gender marking in their pronominal paradigms. From the list of forms that is presented by McElhanon (1975:548), it appears that the independent pronouns from the Papuan language Rawa is of the same paradigmatic structure as the pronouns from Aleut in (8.48). However, the presented first person plural form *nâre* is probably a mistake. The comparison with the possessive suffixes suggest that it should be *nâne* (McElhanon, 1975:550).

8.8 Structural analysis

8.8.1 Preamble

In the preceding survey of the paradigmatic variation, 32 different kinds of paradigms are presented. Some of these paradigmatic structures are only attested in one isolated case; others are found widespread through the world's linguistic variation. In this section, the structural characteristics of the variation are analysed. For this analysis, the individual elements in the paradigm are the locus of investigation, not the paradigmatic structure as a whole (cf chapter 5). The internal structure of the paradigms is taken apart. In fact, a paradigm consists of multiple oppositions between the various elements. The oppositions attested are counted in a sample of pronominal paradigms and there will turn out to be all kinds of interdependencies between various oppositions. The result will be a hierarchy of oppositions, that will be called the 'Dual Explicitness Hierarchy.'

The analysis is carried out on the basis of a sample of 83 cases. These cases are selected from the paradigmatic variation as presented in this chapter. The sample has been constructed as a diversity sample, including every paradigmatic variant that has been attested. In contrast, the influence of structures that are attested in widespread genetic families has been reduced. Maximally two cases per family are counted, interpreting 'family' rather loosely as a close genetic grouping that shows identical paradigmatic structures throughout the genetic group. Moreover, the special case of Palaung (see section 8.5.5) and the various paradigmatic structures from Kapau (see section 8.7) have not been included in the sample. The composition of the sample is described in detail in Appendix C. These criteria used in constructing the sample have a few consequences for the interpretation of the results. First, the resulting frequency of each particular paradigmatic structure is not an accurate representation of the world-wide proportion. Common structures will be somewhat less common than they actually are and the frequency of rare structures will be slightly overestimated. Second, the resulting sample emphasizes the variability of the structure of human language. Every single aberrant case has been included in the sample, but the number of common paradigms has been restricted. A consequence of this approach is that generalisations become less strong. However, a good theory about human language should be able to account for the wide variety of possibilities attested, not only for the common strategies. The following characteristics have been considered in this sample. First, the structure of the first person complex is included. Second, the presence or absence of vertical homophony is included. Vertical homophony is defined as an overlap between the marking of the second and third person dual, or between either the second or third person dual and the first person complex. Exactly which kind of vertical homophony is attested in the paradigm is also included in the set of characteristics. Both the structure of the first person complex and vertical homophony will be discussed in section 8.8.2. Third, the presence or absence of horizontal homophony is included in the investigation. Horizontal homophony is defined as an overlap between the marking of dual and non-dual categories. The distribution of the various

kinds of horizontal homophony is discussed in section 8.8.3. The results of these analyses are summarised in section 8.8.4.³⁵

8.8.2 Vertical homophony and the Explicitness Hierarchy

Vertical homophony is defined as a cluster of referential possibilities in the restricted group column that includes at least the second or third person. An example of vertical homophony is presented in (8.49). This paradigm is the past/future marking form the Papuan language Yagaria (Renck, 1975:90-96). A vertical homophony between the second and third person dual is attested in the form of the suffix *...-a?*.

(8.49) YAGARIA

		<i>group</i>	<i>restricted group</i>	
		...-un	...-u?	1+2
1	...-u			1+2+3
2	...-an			1+3
3	...-i	...-a	...-a?	2+3
				3+3

Very few cases of vertical homophony have been attested in paradigms that distinguish between inclusive and exclusive (see section 8.5.3). This claim is reformulated quantitatively for the present sample of 83 cases in Table 8.1. The observation holds, although the significance is not as strong as the correlations that were found in Chapter 5. The distribution can be summarised by an implication:

(8.50) inclusive/exclusive opposition \rightarrow no vertical homophony

		<i>Inclusive/Exclusive opposition in dual</i>		
		-	+	
<i>Vertical Homophony in dual</i>	-	25	42	67
	+	12	4	16
		37	46	83

Table 8.1: Implication ‘inclusive/exclusive opposition \rightarrow no vertical homophony’
(Fisher’s exact $p = .011$, $\phi = -.299$)

³⁵ No counts of the proportion of inflectional versus independently marked paradigms have been made (as had been done for the paradigms without duals, see section 5.8). By collecting the examples for this chapter, I noticed that comparable paradigms were often found both as independent pronouns and as verbal inflection. In many of those cases, the independent pronoun paradigm most clearly showed the paradigmatic structure, and I decided to take the independent pronouns as example for the description of the diversity. Only when the inflection had a different structure, have I included the inflectional paradigm. Because of this rather ad-hoc decision, I do not expect the proportions of inflectional versus independently marked paradigms as presented in this chapter to be representative of the linguistic structure. Future research will have to attend to this dimension more carefully.

The kind of vertical homophony that is attested is shown in Table 8.2. There is a clear predominance of the homophony between second and third person dual over all other possible kinds of homophony. This distribution is rather different from the occurrence of vertical homophony as attested in paradigms without dual forms (see section 5.3.3). In those paradigms, the homophony between second and third person was also attested but it was the least frequent kind of vertical homophony. The large number of paradigms with a second/third homophony as attested in this chapter is mainly due to languages from the southeastern part of New Guinea, mainly from the languages of the East New Guinea Highlands. The second/third homophony is a clear areal characteristic, found in many languages in this area (Foley, 1986:72). Still, the genetic variety of languages that show this homophony is large; the second/third homophony in the dual is a pattern to reckon with.

<i>kind of vertical homophony</i>	<i>number of cases</i>
second + third person dual	13
first + third person dual	1
Inclusive + third person dual	1
Exclusive + second person dual	1

Table 8.2: Various kinds of vertical homophony of dual forms

The paradigmatic structure of the dual morphemes also shows an Explicitness Hierarchy, analogously to the structure of paradigms without a dual (see section 5.5). This DUAL EXPLICITNESS HIERARCHY is in fact an extension of the implication as discussed above in (8.50). For the formulation of the Dual Explicitness Hierarchy the structure of the restricted group column is scrutinised. The structure of the rest of the paradigm is disregarded temporarily. Three characteristics of the restricted group column are relevant:

- Split inclusive:** there is an opposition between the minimal inclusive (1+2) and the augmented inclusive (1+2+3);
- Split dual-‘we’:** there are two different forms for ‘we’ in the restricted group column: a difference is marked between inclusive and exclusive dual;
- Split dual:** there is no vertical homophony in the dual, ie the second and third person dual are different from each other and different from the dual first person complex.

The distribution of these characteristics in the sample is shown in Table 8.3. The majority of the attested patterns follows an hierarchical ordering of the three characteristics: split inclusive > split dual-‘we’ > split dual. This ordering is called the Dual Explicitness Hierarchy.³⁶

³⁶ The middle range of the hierarchy appears to be overrepresented. However, this is a side-effect of the sampling method. The middle range includes all paradigmatic structure with horizontal homophony. Because I wanted to show the complete paradigmatic variation, I have included all different cases with horizontal homophony (see the next section).

	<i>common</i>				<i>rare</i>			
<i>Split inclusive</i>	+	-	-	-	-	+	+	+
<i>Split dual- 'we'</i>	+	+	-	-	+	-	+	-
<i>Split dual</i>	+	+	+	-	-	+	-	-
Number of cases	13	25	28	13	2	1	1	0
	79 cases (95%)				4 cases (5%)			

Table 8.3: The Dual Explicitness Hierarchy

The Dual Explicitness Hierarchy is schematically presented by the paradigmatic structures in Figure 8.3. The columns in the figure represent the structure of the restricted group column. To the left, all categories are distinguished by separate morphemes, losing more and more explicitness going to the right.

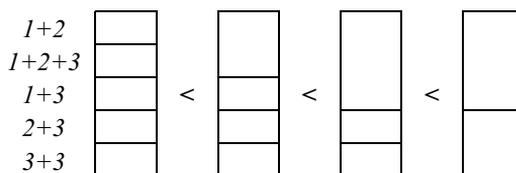


Figure 8.3: Schematic structure of the 'restricted group' column in the different stages of the Explicitness Hierarchy

This Dual Explicitness Hierarchy runs parallel to the Explicitness Hierarchy that was found for paradigms without dual forms. In the next chapter, these two hierarchies will be shown to be interconnected by various examples of cognate paradigms.

8.8.3 Horizontal homophony

Next to vertical homophony, horizontal homophony is the other structural characteristic of a pronominal paradigm to be scrutinised here. Horizontal homophony is characterised by one morpheme that is used for groups of referents with different numbers. In the present context, the important question is whether the dual reference is identical to the reference for any other number category or not. An example of horizontal homophony is presented in (8.51). What is shown are the independent pronouns from Sedang, a Mon-Khmer language. This paradigm has a horizontal homophony in the second person dual (Smith 1979: 80).

(8.51) SEDANG

		<i>group</i>	<i>restricted group</i>	
		pin	pá	1+2
		ngin	má	1+2+3
1	á	pó		1+3
2	eh			2+3
3	gá	vai	préi	3+3

I have tabulated the frequencies of the possible combinations of horizontal homophony within a paradigm in Table 8.4. A minus in the table means that there is no horizontal homophony between the dual and the plural, which can also be interpreted as meaning that there exists a specialised dual form. A plus indicates that there is horizontal homophony, which can also be interpreted as meaning that there is no specialised dual form. Two columns in the table below are marked as ‘no dual’ because these combinations of plusses result in a paradigm that does not mark number anymore.³⁷

<i>Inclusive</i>	–	–	–	–	+	+	–	+	+	–	–	+	–	+	+	+
<i>Exclusive</i>	–	–	–	+	+	+	–	–	+	+	+	+	+	–	–	–
<i>Second person</i>	–	–	+	+	+	+	+	–	–	–	+	–	–	+	–	+
<i>Third person</i>	–	+	+	+	+	–	–	–	–	–	–	+	+	–	+	+
Number of cases	59	5	4	(no dual)		4	4	3	1	1	1	1	0	0	0	0

Table 8.4: Horizontal homophony between dual and non-dual forms

It is tempting to try and see a hierarchy in the frequencies, but the apparent hierarchy is not very relevant. In fact, the far majority of cases has simply no horizontal homophony at all (59 cases, which is 71% of the sample). The next two most frequent cases fall neatly into the apparent hierarchy (respectively 5 and 4 cases, which is 11% of the sample), but there are still another 15 cases left (18% of the sample) that do not fit into the hierarchy. The large majority of 15 combinations of horizontal homophony have at least a specialised dual in the third person. This can be explained by noting the strong correlation between duals in the 3rd person and nominal dual marking (Plank, 1996:133).

The claim by Humboldt and Plank that duals are preferably found in the first person (see the introduction of this chapter) is only true if the minimal inclusive forms are counted as duals. The minimal inclusive the category for the pair of speaker and addressee. This ‘natural pair’ is dual by definition, but I have argued elsewhere that this category is better interpreted as a special group category, and not as a dual (see sections 3.6.5 and 4.7.1). When the ‘natural pair’ is not included among the duals (as I have done), then almost all paradigms take all or nothing.³⁸ The best generalisation over the presented distribution is that dual marking is preferably found across the board of the paradigms, or not at all. Intermediate cases with special combinations of horizontal homophony are rare.

³⁷ The column with a minus for inclusive and plusses for all other persons can be interpreted as referring to the Maranao-type paradigm with a minimal/augmented pattern (see section 4.7.2). The column with all plusses corresponds to the paradigms without dual marking. These paradigms have been discussed extensively in Part II of the present work.

³⁸ There are 24 cases in the sample of 83 paradigms that do not have a specialised dual for all persons (specialised dual forms are marked by a minus in Table 8.4). The claim by Humboldt and Plank is restricted to this set of cases. Out of these 24 cases, 18 paradigms (22% of 83 paradigms) have some kind of dual ‘we’, 11 paradigms (13% of 83 paradigms) have a second person dual and 15 paradigms (18% of 83 paradigms) have a third person dual. Indeed, dual ‘we’ is the most frequent (as claimed by Humboldt and Plank), but the preference is not really impressive.

Finally, the presence or absence of horizontal homophony is crossed with the Dual Explicitness Hierarchy. Horizontal homophony is almost exclusively attested in the middle range of the Dual Explicitness Hierarchy, as is shown in Table 8.5. This restriction can be explained by two constraints on the structure of a pronominal paradigm (cf section 5.6.3). First, vertical homophony and horizontal homophony are preferably not combined into one paradigmatic structure. Second, the opposition between minimal inclusive and augmented inclusive is the last addition to a pronominal paradigm. This opposition is only marked in a paradigm after all other possible opposition have already been expressed.

		<i>Dual Explicitness Hierarchy</i>				
		+	-	-	-	others
<i>Horizontal Homophony</i>	-	13	14	17	12	3
	+	0	11	11	1	1

Table 8.5: Correlation between the Dual Explicitness Hierarchy and the presence of horizontal homophony. Horizontal homophony is almost exclusively found in the middle range of the Dual Explicitness Hierarchy ($\chi^2 = 12.40$, $DF = 4$, $p < .015$).

8.8.4 Summary

The investigation of the structural characteristics of the dual has resulted in some interesting generalisations. First, it was noted that there is an inverse correlation between an inclusive-exclusive opposition in the dual and the occurrence of vertical homophony. This result was already foreshadowed on qualitative grounds in section 8.5.3. In this section, it turned out that also quantitatively a (moderately) significant inverse correlation is found in the sample. Second, this correlation has been shown to be part of a larger Dual Explicitness Hierarchy. This hierarchy describes the strong tendency to explicate the dual categories in a strict order: split dual > split dual-‘we’ > split inclusive. The higher up the hierarchy, the more explicit the marking of the dual. This hierarchy is the dual analogue to the Explicitness Hierarchy that was found in section 5.5. Third, the distribution of the various forms of horizontal homophony did not show much of interest. There is a faint feeling of a hierarchy, but across the board the main representatives of the hierarchy are the endpoints (either there is a complete dual set or there is no dual at all). The intermediate stages of a putative ‘dual’ hierarchy 1>2>3 are only sparsely represented and the counterexamples to the hierarchy are numerous. Finally, horizontal homophony is almost exclusively attested in the middle range of the Dual Explicitness Hierarchy. This correlation can be explained by positing two constraints on paradigmatic structure. Horizontal and vertical homophony are not to be combined into one paradigm, and the minimal/augmented opposition is only overtly marked after all other possible oppositions have been expressed.

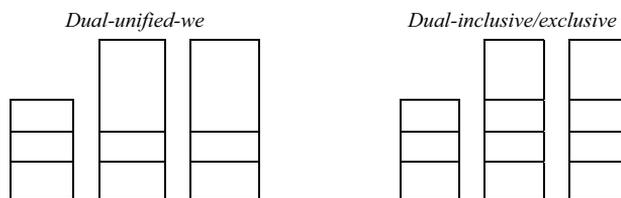


Figure 8.5: The two most frequent paradigmatic structures

The second set of paradigms on the frequency continuum is formed by the two paradigmatic structures shown in Figure 8.6. Structurally, both patterns have a distinction between minimal and augmented inclusive. Areally, both patterns are almost exclusively found in northern Australia. The most well known of these patterns is the UNIT-AUGMENTED paradigm (see section 8.6). A variant of the unit-augmented paradigm is the PARTIAL-UNIT-AUGMENTED paradigm. The only difference between the two is that the ‘inclusive trial’ is not attested in the partial-unit-augmented pattern. The resulting partial-unit-augmented paradigm is functionally equivalent to the dual-inclusive/exclusive type (cf Figure 8.5 above). However, the morphological structure of the partial-unit-augmented paradigms indicates a connection to the unit-augmented paradigm (see section 8.5.4). The partial-unit-augmented paradigm is kept separate, because it will turn out to play an important role in the next chapter, mediating diachronically between other major patterns.

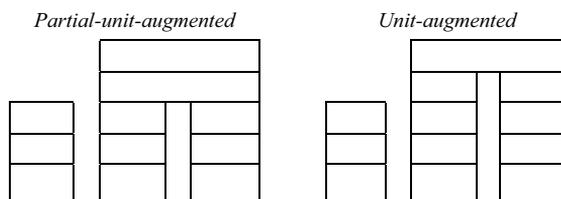


Figure 8.6: Major paradigmatic structures with a minimal/augmented opposition

The third kind of paradigmatic structure highlighted here is a pattern with three different forms for ‘we’. There are many paradigmatic variants attested for the marking of the three forms for ‘we’. All these paradigmatic possibilities have been combined under the heading of DUAL-3WE (see section 8.4). The referential values of these three different forms for ‘we’ are highly variable. At least five different patterns were distinguished and there might be even more. The large majority of dual-3we paradigms is found in northern Australia and in southeastern New Guinea. Another characteristic shared by all these patterns is that they are idiosyncrasies within their close family. The paradigms with three forms for ‘we’ are never a general characteristic of a whole family. In the next chapter, it will be shown that dual-3we patterns are an intermediate stage in between other paradigmatic structures, which explains their highly variable nature. Only the most frequent of these patterns is shown in Figure 8.7.

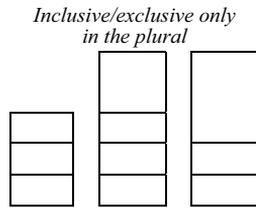


Figure 8.7: The most frequent of the various dual-3we paradigmatic structures

The final set of relatively common paradigmatic is formed by the two paradigmatic structures that are shown in Figure 8.8. These patterns both have a vertical homophony between the second and the third person. The KALAM-TYPE shows only a vertical homophony in the dual; the YAGARIA-TYPE shows vertical homophony both in the dual and in the plural. Vertical homophony between addressee and other is rather frequent in paradigms with dual marking. Interestingly, in paradigms without dual marking, this kind of vertical homophony is much less common (relative to other kinds of vertical homophony). In line with the geographical distribution of the dual-3we paradigms, all examples of the two paradigmatic structures as shown in Figure 8.8 are from the southeastern part of New Guinea, with a strong emphasis on the East New Guinea Highlands (see section 8.3.3).

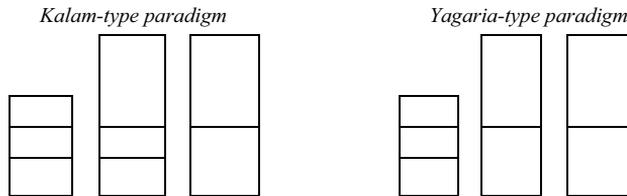


Figure 8.8: Major paradigmatic structures with a vertical homophony

These various paradigmatic structures are not accidental phenomena. The fact that exactly these patterns are attested more often than others is the result of a multifaceted force-field that shapes linguistic structure. As a first impetus towards unravelling the complex of factors governing the structure of pronominal paradigms, the complete variation of all 32 different paradigmatic structures has been analysed on structural characteristics within the paradigm (see section 8.8). A strong typological constraint on the organisation of the dual marking in pronominal paradigms has been formulated in the form of the Dual Explicitness Hierarchy. The Dual Explicitness Hierarchy describes the tendency for dual marking to be more or less explicit along a line of four stages. These four stages are illustrated by their most common representatives in the upper row of Figure 8.9. The least explicit form of dual marking only distinguishes a dual ‘we’ from a dual ‘non-we’ (the YAGARIA-TYPE paradigm). In the next stage, the dual ‘non-we’ is split into a dual second person and a dual third person (the DUAL-UNIFIED-WE paradigm). In the third stage, the marking of dual ‘we’ is split into an exclusive and an inclusive form (the DUAL-INCLUSIVE/EXCLUSIVE paradigm). Finally, an opposition between minimal and augmented inclusive is added to the marking of explicitness in a paradigm (the UNIT-AUGMENTED paradigm).

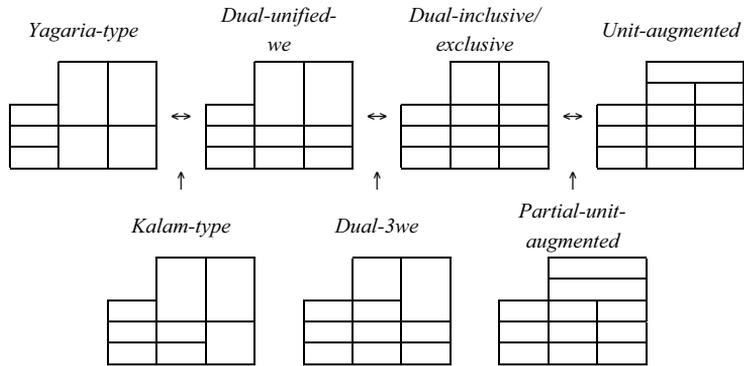


Figure 8.9: The Dual Explicitness Hierarchy and some intermediate forms

The paradigms in the upper row of Figure 8.9 all show a symmetry between the dual and the plural column. The other frequently attested paradigmatic structures are shown in the second row of the figure. They can be interpreted as intermediate forms between the major stages of the Dual Explicitness Hierarchy. Together, all these patterns form a web of interconnected paradigmatic structures. The central part of this web is formed by the two most frequently attested and geographically most widespread paradigmatic structures (the dual-unified-we and the dual-inclusive/exclusive paradigm). The various paradigms around these two central figures are less common. Moreover, these patterns are only found in a very restricted geographical area. The paradigms on the fringes of Figure 8.9 are almost exclusively found in northwestern Australia and southeastern New Guinea.

In the next chapter, I will investigate how this web of interconnected paradigmatic structures fares within the confines of the variation between closely related languages. Miscellaneous cases of cognate paradigms will show that this hierarchy is not only a structural characteristic of human language, but that it also has diachronic repercussions.

Chapter 9

Cognate paradigms revisited

Connecting the dual through time and space

9.1 Introduction

A pronominal paradigm is not an isolated, unchangeable structure. The structure of a pronominal paradigm is highly variable through time and space and the variation in paradigmatic structure, even between closely related languages, is remarkable. Individually, pronominal elements are rather stable, but complete pronominal paradigms change easily. This variability might make one weary of a typological generalisation or of a diachronic analysis of the paradigmatic structure because both types of investigation need an object with some stability to yield results. However, both kinds of analyses together – typology and diachrony combined – can overcome this variability. The diachronic study of human language starts from small variations between close relatives to develop generalisations about historical change. Typology, on the other hand, abstracts away from the small variations and focuses on the major breakpoints in the general structure of human language. In the previous chapter, I have outlined a typology of paradigmatic variation of dual marking in the pronominal domain. In this chapter, I will bring together this typology with a special type of comparative analysis, which will result in a cognitive map of interrelated paradigms, approximating the dynamics of paradigmatic structure through time and space.

9.2 Method of comparison

Before I proceed, a few words on this special kind of comparative method are in order. In most typological research, variation between closely related languages is disregarded. However, this seems to be a waste of information, at least for the present case. It is highly instructive to analyse the variation between close relatives because large differences are attested within the confines of the narrow linguistic family. Closely related languages normally show small differences in the paradigmatic structure of their pronominal elements. Sometimes, these small differences are typologically salient divisions between major paradigmatic structures. Yet, the morphemes themselves, of which the paradigms are built, are still clearly recognisable as cognates. Such pairs of paradigms, which consist of different paradigmatic structures, yet built from clearly cognate morphemes, are called ‘cognate paradigms’. Numerous cases will be presented in this chapter. The comparison of cognate paradigms comes close to a historical investigation, but it should not be considered identical to it. A true

diachronic investigation into paradigmatic structure will need to look much more closely into the minutiae of variation. In this context, I will look only at cases where history has led to specific differences in the paradigmatic structure. The cognate paradigms presented differ only in one morpheme – or in very few semantically related morphemes. I have searched for cases of paradigms that are extremely closely related, so various intermediate steps seem unwarranted to explain the change from one paradigm to the other. Also, for the cases presented, it seems unnecessary to propose that the cognate paradigms go back to a third, unattested, paradigmatic structure.¹ As a result, the change from one paradigm into the other could be either an addition or a loss of a few morphemes. Which of the two directions accurately describes the historical change is left undecided in most cases. It is often difficult to establish a concise historical development that leads to the present situation only on the basis of the different end results. Consequently, I will propose historical developments without the direction of change for most of the cases to be discussed below. A true historical reconstruction of paradigmatic structure is left for another occasion. The present goal is to present a rough outline of the paths along which pronominal paradigms can change, and suggest some restrictions on the possible diachronic modifications of paradigmatic structure.²

Only a selected set of cognate paradigms will be discussed. The focus will be on the interrelation between the major paradigmatic patterns that were distinguished in the previous chapter. These major patterns are repeated schematically in Figure 9.1. In each of these patterns, the three columns represent ‘singular’, ‘group’ and ‘restricted group’ respectively. These major patterns are ordered along the Dual Explicitness Hierarchy as formulated in the previous chapter. This hierarchy describes a strong constraint on the order in which particular oppositions are marked in a pronominal paradigm. As a hypothesis, this typological hierarchy is interpreted as a pathway for diachronic change. I will use the examples of cognate paradigms to test this hypothesis. If the hierarchy stands diachronic interpretation, I expect to find examples of cognate paradigms linking two paradigms that are adjacent on the hierarchy. In general, the hypothesis will fare rather well. The examples of cognate paradigms that will be discussed in this chapter support the Dual Explicitness Hierarchy to a large extent.

¹ Of course, all cases of cognate paradigms go back to a third, unattested, paradigmatic structure: the proto-structure. Both extant paradigms – now called ‘cognate’ – are the result of different changes from a historical predecessor. Literally speaking, the change is never from one cognate to another, but for both languages from the proto-structure to the extant structures. However, the cases that will be presented here seem to be situations where one of the extant structures did not (or hardly) change since the proto-structure that united the two cognates. (At least, the paradigmatic structure of this ‘conservative’ paradigm did not change. Phonologically, the ‘conservative’ paradigm might be highly ‘progressive’; cf the case of Uradihi in (9.30) below.) When I talk of a change from one extant structure to another, this should be taken as shorthand for a situation where one of the two structure is much closer to the paradigmatic structure of the proto-language than the other.

² See section 6.2 for a more extensive discussion of the methodological aspects of the comparison of cognate paradigms.

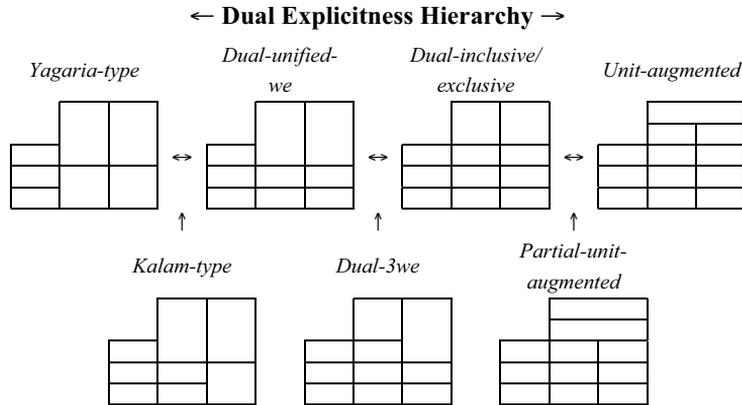


Figure 9.1: Major paradigmatic types with 'restricted group' morphemes ordered along the Dual Explicitness Hierarchy

The complete typology of the paradigmatic structure of dual forms was described in the previous chapter. Only a few comments are repeated here. First, only the MAJOR patterns are presented. This is an abstraction from the large variety of paradigmatic structures that were attested. However, none of the other patterns is attested commonly enough to grant it the status of a 'major' type. Second, the DUAL-3WE pattern exists in many different variants. Only one of the five different paradigmatic structures that were attested is shown in Figure 9.1 (cf section 8.4). The variability of this type is mirrored in the many cases of cognate paradigms that link the dual-3we pattern to the various other paradigmatic structures. Section 9.5 will be devoted to the discussion of the connections between the dual-3we pattern and other patterns. It turns out that the dual-3we pattern is a mediator between the other patterns with dual marking. Third, two of the various types that are distinguished in Figure 9.1 are FUNCTIONALLY identical. The dual-inclusive/exclusive pattern and the partial-unit-augmented pattern mark exactly the same referential categories. The only reason to distinguish between the two is a difference in their morphological structure (cf section 8.5.4). At first, these two will be kept apart in this chapter because they show different connections to other paradigms. However, the two will be equated for the final version of paradigmatic interconnectivity in section 9.6.

The major types from Figure 9.1 will not only be linked to each other, they will also be linked to the major paradigmatic structures without a dual. The patterns without a dual that will be considered in this chapter are presented in the upper row of Figure 9.2. These paradigmatic structures are ordered along the Explicitness Hierarchy, which runs parallel to the Dual Explicitness Hierarchy. As a diachronic interpretation, I hypothesise that the two hierarchies are connected to each other. This means that paradigms are expected to change along the lines as shown in Figure 9.2. If this hypothesis is any good, then cognate paradigms are expected to turn up linking adjacent paradigmatic structures in the figure.

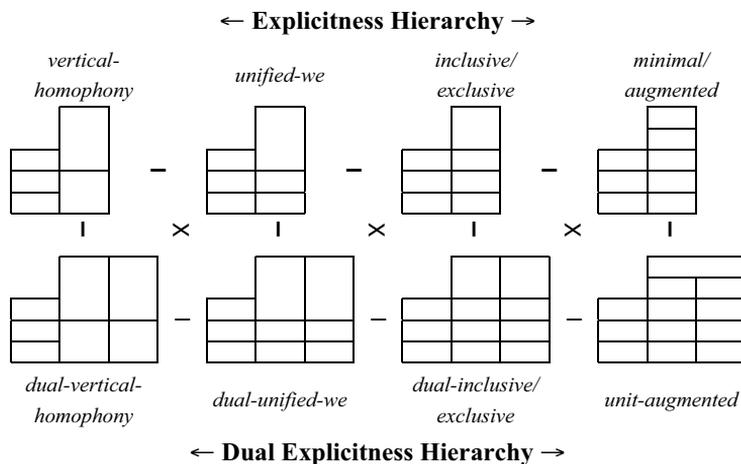


Figure 9.2: Hypothesised connections between the Explicitness Hierarchy and the Dual Explicitness Hierarchy

In this chapter, I will present numerous cases of cognate paradigms, which will approximately confirm the hypothesised connections from Figure 9.2. After each set of examples, I will give some arguments why specific links are not attested. These arguments are not final, and could be falsified by a good example showing unambiguously that a particular connection exist. However, as long as I do not have any cases that point towards a direct connection, I assume that these links do not exist. The connections attested between the various patterns will be discussed in three sections. First, the examples involving the typical DUAL patterns are discussed (the dual-vertical-homophony, the dual-unified-we and the dual-inclusive/exclusive paradigms). These paradigms will be linked to each other and to the corresponding non-dual patterns in section 9.3. Second, the MINIMAL/AUGMENTED opposition in all its variations (the minimal/augmented, the partial-unit-augmented and the unit-augmented paradigms) will be shown to be connected in section 9.4. These links are added to the emerging web of interconnected paradigmatic types. Finally, the DUAL-3WE pattern and its multiple connections are discussed in section 9.5. Abstracting somewhat from the quirky ways of language, this complete web of interrelated paradigmatic structures will be reduced to a more conveniently arranged cognitive map in section 9.6. This cognitive map will be extended with the connections that were already established in chapter 6. The complete map will be analysed in detail in section 9.7. Finally, this chapter will be summarised in section 9.8.

9.3 Linking the major dual paradigms

9.3.1 Preamble

In this first part of the discussion of cognate paradigms, I will present examples of cognate paradigms involving the typical dual paradigms. They will be linked to each other and to the corresponding paradigms without a dual. The hypothesised intercon-

nections between the various paradigms are repeated in Figure 9.3. The three structures that will be discussed in this section are marked in grey.

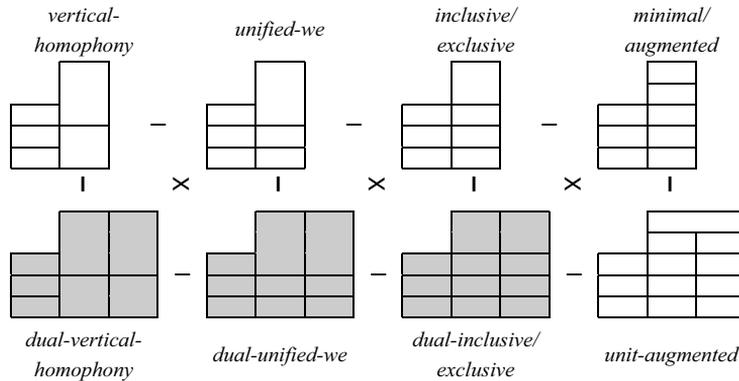


Figure 9.3: The major dual paradigms within the hypothesised paradigmatic connections

9.3.2 Cognate paradigms

Examples will be discussed from Finisterre-Huon, Pama-Nyungan, Uralic, Kiranti and Miwok; a typologically, genetically and areally highly diverse set of examples. The first two cases are from languages that are spoken in the Pacific area. The Finisterre-Huon family is a group of related Papuan languages from New Guinea and the Pama-Nyungan family is a large group of related languages from Australia. The next two cases come from Eurasia. The Uralic family is a group of far-flung languages spoken throughout Russia and the Kiranti family is a sub-group of the Tibeto-Burman stock, spoken in Nepal. Finally, the Miwok languages are spoken in California.

Finisterre-Huon (New Guinea)

The independent pronouns from the Finisterre-Huon language Wantoat are shown in (9.1). These pronouns have the typical New Guinea Highlands structure, showing a homophony between the second and the third person non-singular (McElhanon, 1975:548). Wantoat is a language from the Finisterre-range. In the independent pronouns from Kewieng, another language from the Finisterre range, the same structure is found. The somewhat loosely related languages from the Finisterre range belong quite probably to the same stock as the closely-knit Huon family (Foley, 1986:244). The independent pronouns from the Huon language Nabak are shown in (9.2). These pronouns are almost identical to the Wantoat pronouns, except for the opposition between the second and the third person in the non-singular. The same paradigmatic structure as found in Nabak is also found in other Huon languages, such as Selepet, Ono, Kube, Kâte and Kovai (McElhanon, 1975:548).³

³ The third person pronouns in Nabak show signs of recent addition. The singular morpheme *ek* is compounded with the second person non-singular morphemes to form third person non-singular forms. Probably, the Nabak structure is diachronically later than the Wantoat structure.

(9.1) WANTOAT

		<i>group</i>	<i>restricted group</i>	
		nin	nit	1+2
		gin	git	1+2+3
1	nâ			1+3
2	gâ			2+3
3	an			3+3

(9.2) NABAK

		<i>group</i>	<i>restricted group</i>	
		nen	net	1+2
		in	it	1+2+3
1	nâ			1+3
2	gâ			2+3
3	ek	ekŋen	ekget	3+3

Pama-Nyungan (Australia)

Another obvious link between various dual patterns is the connection between the dual-unified-we paradigm and the dual-inclusive/exclusive paradigm. Starting from the dual-unified-we paradigm, only an inclusive-exclusive distinction has to be added (both in the group and in the restricted group morphemes) to arrive at a dual-inclusive/exclusive pattern. In (9.3), the independent pronouns from the Pama-Nyungan language Warrgamay are shown (Dixon, 1981:40). This is a typical paradigm among the Pama-Nyungan languages. In 9.4), the independent pronouns from the neighbouring Pama-Nyungan language Nyawaygi are shown (Dixon, 1983:463-467).

Almost the same pronouns are attested; the only difference is the appearance of separate exclusive forms. These exclusive forms are derived from the inclusive forms by a suffix *...-liju*. The inclusive forms in Nyawaygi are the same as the unified-we morphemes from Warrgamay in (9.3). The dual-inclusive/exclusive paradigm from Nyawaygi is derived from a dual-unified paradigm as found in Warrgamay. This development is common among Australian languages:

‘More than half the languages with a singular/dual/plural pronominal system show an inclusive/exclusive distinction, but there is no regularity to the distribution – languages of both types are found in every quarter of the continent. And whereas the four forms we reconstructed for 1 dual, 1 plural, 2 dual and 2 plural cannot be further analysed, it is nearly always possible to provide some analysis of inclusive/exclusive forms. ... This suggest most strongly that an inclusive/exclusive distinction should not be attributed to pA [proto-Australia] but has evolved rather recently in a number of scattered groups of modern languages.’ (Dixon, 1980:335-336)

(9.3) WARRGAMAY

		<i>group</i>	<i>restricted group</i>	
		ɲana	ɲali	1+2
1	ɲayba			1+2+3
2	ɲinba	ɲura	ɲubala	1+3
3	ɲaŋa	ɖana	bula	2+3
				3+3

(9.4) NYAWAYGI

		<i>group</i>	<i>restricted group</i>	
		ɲana	ɲali	1+2
1	ɲayba	ɲanaliŋu	ɲaliliŋu	1+2+3
2	ɲinba	ɲura	ɲubula	1+3
3	ɲaŋga	ɖana	bula	2+3
				3+3

Uralic (Russia)

The next two cases of cognate paradigms will highlight the connections between the paradigms with a dual and the paradigms without a dual. A straightforward link is the connection between the unified-we and the dual-unified-we pattern. The only difference between these two patterns is a set of dual markers that is added to the set of group markers. This link is exemplified with independent pronouns from two Uralic languages; Udmurt (Csúcs, 1998:287) and Nganasan (Helimski, 1998:501) in (9.5) and 9.6). The Nganasan paradigm has added dual pronouns by inserting a new set of group pronouns and reanalysing the old group pronouns as duals.⁴

(9.5) UDMURT

		mi	1+2
1	mon		1+2+3
2	ton	ti	1+3
3	so	soos	2+3
			3+3

(9.6) NGANASAN

		<i>group</i>	<i>restricted group</i>	
		mǎŋ	mi	1+2
1	mənə			1+2+3
2	tənə	tǐŋ	ti	1+3
3	sǐtǐ	sǐtǐŋ	sǐtǐ	2+3
				3+3

⁴ This is a case of morphological markedness reversal, cf Ika and Damana in section 7.3

Kiranti (Nepal)

The next cognate paradigms are found among the independent pronouns of the Kiranti languages (part of Sino-Tibetan), spoken in Nepal. The following languages have been selected out of the large variety of Kiranti languages because the pronominal forms are almost identical. No argumentation is needed to show that these paradigms are cognate.⁵ Still, there are some interesting paradigmatic differences between the languages. First, the pronouns from Thulung are shown in (9.7). This paradigm has an inclusive/exclusive opposition, but no dual forms (Bauman, 1975: 126-127).

(9.7) THULUNG

		goi	1+2
		goku	1+2+3
1	go	goku	1+3
2	gana	gani	2+3
3	(not given in source)		3+3

Second, the pronouns from Bahing are shown in (9.8). These pronouns have a dual column, which results in a dual-inclusive/exclusive paradigm (Bauman, 1975:267). The dual column is clearly marked relative to the group column by a suffix *...-si*. However, it is not unequivocally clear that this represents a diachronically later addition because the suffix *...-si* appears infix in the 1+3 form *go:su:ku*. These paradigms from Thulung and Bahing are clearly cognate, but the historical direction of the changes leading to the present variety remains to be investigated by specialists. For the present purpose, it suffices to note that the dual-inclusive/exclusive paradigm is closely related to the inclusive/exclusive pattern.

(9.8) BAHING

		<i>group</i>	<i>restricted group</i>	
		go:i	go:si	1+2
		go:ku	go:su:ku	1+2+3
1	go	go:ku	go:su:ku	1+3
2	ga	gani	gasi	2+3
3	harem	haremdau	haremdausi	3+3

Miwok (USA)

A striking case of cognate paradigmatic variation is found among the Miwok languages, spoken in California. This case was already noted by Freeland (1947:35), Callaghan (1974:384-385) and Greenberg (1988:9-11). The differences between the independent pronouns of the various Miwok languages are small and it seems to be within reach to reconstruct the complete history of the paradigmatic changes in the Miwok pronoun paradigm, though this task falls outside the scope of the present

⁵ In contrast to the cases presented, the pronouns from other Kiranti languages are in need of extensive historical-comparative backing before it can be maintained that they are related, cf Bauman (1975:123-142), van Driem (1987:25-28) and Ebert (1997:43) for independent pronouns from some other Kiranti languages that are not immediately clear as cognate paradigms.

study. Clearly, a direct change from an extant paradigm of one of the Miwok languages into another, as proposed by Greenberg (1988:9-11), does not do justice to the intricacies of the variation attested. Still, there are clear cognate paradigms with different paradigmatic structures among the Miwok languages. These indicate that there is a link between the inclusive/exclusive, the minimal/augmented and the dual-unified-we paradigms. The orthography of the Miwok pronouns is taken from Callaghan (1974).

The differences in paradigmatic structures between the various the Miwok languages centre around two linguistic elements, reconstructed as **ʔičy-* and **ʔotĩ-*. The morpheme **ʔotĩ-* is a numeral ‘two’. The history of **ʔičy-* is unknown (Callaghan, 1974:386). The reflexes of these proto-morphemes will be found in various functions in the different Miwok languages. The first variant is the paradigm of the independent pronouns from Southern Sierra Miwok, shown in (9.9). This paradigm has an opposition between 1+2 and 1+2+3, forming a minimal/augmented paradigm. Both inclusive forms are made on the basis of the proto-Miwok **ʔotĩ-*. The second part of these inclusive pronouns is, on the one hand, the first person plural suffix *...-me* for 1+2 and, on the other hand, the first person inclusive plural possessive suffix *...-c'i* for 1+2+3 (Broadbent, 1964:93; Callaghan, 1974:384). This possessive suffix *...-c'i* is possibly historically related to the proto-Miwok **ʔičy-*.

(9.9) SOUTHERN SIERRA MIWOK

		ʔotĩme-	1+2
		ʔotĩ-c'i-	1+2+3
1	kan'i-	mah'i-	1+3
2	mí'ni-	mi-ko-	2+3
3	ʔis'ak-	ʔi-k'o-	3+3

A different pattern is found in Bodega Miwok, shown in (9.10). In this paradigm, the pronoun *ʔóc'i* is the same form as the proto Miwok numeral **ʔotĩ-* by a regular sound change (Callaghan, 1974:385). In this case, the numeral ‘two’ has been reinterpreted as a first person dual and duals for the other persons are added analogously. In this way, a dual-unified paradigm has been formed.⁶

(9.10) BODEGA MIWOK

		<i>group</i>	<i>restricted group</i>	
		má'-ko	ʔóc'i	1+2
				1+2+3
1	kán'i			1+3
2	mí'	mí-k'o	mí-k'oʃ	2+3
3	ʔítĩ	ʔi-k'o	ʔi-k'oʃ	3+3

⁶ In Lake Miwok (that has the same paradigm structure as Bodega Miwok) the pronoun *ʔici* is used by one of the informants as an idiosyncratic variant of *ʔoci* (Callaghan, 1965:283). This can be interpreted in different ways. Maybe an erstwhile opposition between *ʔici* and *ʔoci* has merged, or the pronoun *ʔoci* has replaced an older *ʔici*.

The historical developments that resulted in these structures are not completely clear. However, a direct link between the minimal/augmented pattern (9.9) and the dual-unified-we pattern (9.10) seems unlikely. Too many conceptual and morphological changes have to take place for a direct transition. An inclusive-exclusive pattern, as found in Plains Miwok, seems a natural intermediate pattern.

The independent pronouns from Plains Miwok are shown in (9.11). The reconstructed form **ʔič̣y-* is attested as the inclusive first person plural *ʔic̣y-*. The pronoun paradigm thus forms an inclusive-exclusive pattern. However, the status of *ʔic̣y-* as a general inclusive is doubtful. Callaghan (1974:386) analysed the reference of this morpheme as being minimal inclusive only (1+2), but she did not explain in which way the reference to the augmented inclusive (1+2+3) is made. In a later work, she glossed the meaning of *ʔic̣y-* as ‘we inclusive?/you & I’ (Callaghan, 1984:296-297). A comparable restriction of the inclusive *ʔic̣y-* to minimal reference (1+2) is described for a close relative. The independent pronouns of Northern Sierra Miwok, a neighbouring Miwok language, show the same paradigmatic structure as in Plains Miwok. Callaghan (1987:397-398) glosses the morpheme *ʔic̣i-* in Northern Sierra Miwok as ‘thou & I’.⁷ This indicates that the morpheme may be a general inclusive, but that it is prototypically used for the speaker-addressee dyad 1+2. This is represented as the rightmost pattern in (9.11).

(9.11) PLAINS/NORTHERN SIERRA MIWOK

		ʔic̣y-	1+2			ʔic̣y-	1+2
		maṣ'i-	1+2+3			maṣ'i-	1+2+3
1	kaṇ'i-	maṣ'i-	1+3	1	kaṇ'i-	maṣ'i-	1+3
2	mi-	moḳ'o-	2+3	2	mi-	moḳ'o-	2+3
3	ʔiṣy-	ʔi-ḳ'o-	3+3	3	ʔiṣy-	ʔi-ḳ'o-	3+3

The paradigmatic structures from Plains Miwok and Northern Sierra Miwok, as shown in (9.11), are intermediate patterns between the clearly DUAL form in Bodega Miwok and the clearly INCLUSIVE forms from Southern Sierra Miwok. Also geographically, they represent a mediating position. Plains Miwok and Northern Sierra Miwok are found in between Southern Sierra and Bodega Miwok. However, the pronoun **ʔič̣y-* is not indisputable as an intermediate between the other two languages. The major problem is that it can not be reconstructed as an innovation of the numeral **ʔoṭi-*. Still, the Miwok examples suggest two different links: between (9.11) and (9.9) – from inclusive/exclusive to minimal/augmented – and between (9.11) and (9.10) – from inclusive/exclusive to dual-unified-we. In other words, the inclusive/exclusive structure mediates between the other two. Whether the inclusive/exclusive structure as found in Plains Miwok can be identified with the proto-Miwok pronoun paradigm remains an open question.

⁷ However, much earlier, Freeland (1951:30) analyses the morpheme *ʔič̣iʔ* from Northern Sierra Miwok as a general inclusive, encompassing both the minimal (1+2) and the augmented (1+2+3) reference.

9.3.3 Summary

The five cases that were discussed in this section show different connections between paradigmatic structures. These connections are summarised in Figure 9.4. The links on the upper row were already established in chapter 6. Several hypothetically possible connections remain open in the figure. First, connections to the paradigms with vertical homophony (on the far left side of Figure 9.4) are only sparsely attested. This is probably a result of the low frequency of occurrence of these patterns, which makes it less probable that a connection to these paradigms will be found among the world's languages. Second, connections to the paradigms with a minimal/augmented opposition (on the far right side of Figure 9.4) are not shown here, but they do exist. Cases of cognate paradigms involving these patterns will be discussed in the next section.

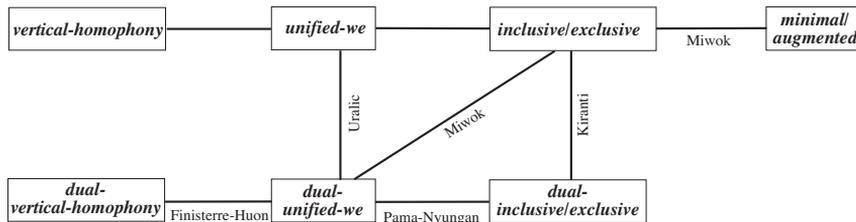


Figure 9.4: First approximation of paradigmatic interconnectivity

The most important missing link among the presented connections is a direct transition between the unified-we paradigm and the dual-inclusive/exclusive paradigm (see Figure 9.5). This direct change seems to be impossible. There are too many paradigmatic changes that have to happen at once for this transition to be instantiated by one change. This transition probably has to be made in several steps, for example, by taking a route over the dual-unified-we pattern, or over the inclusive/exclusive pattern. Of course, there might be cases unknown to me that argue for this direct change, which I argue not to be possible. However, I rest my case here as a falsifiable hypothesis.

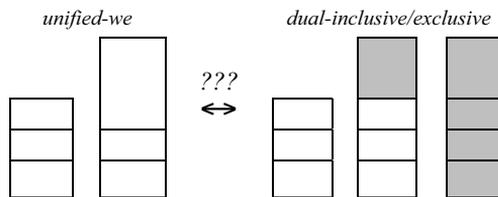


Figure 9.5: Not attested direct link between unified-we and dual-inclusive/exclusive

9.4 Minimal/augmented and its variants

9.4.1 Preamble

In this section, I will present examples of cognate paradigms that link various paradigmatic structures that have an opposition between a minimal and an augmented inclusive. Three different paradigmatic structures with this opposition have been at-

tested frequently. They are schematically shown in Figure 9.6. These three paradigmatic structures are conceptually highly similar and diachronically they seem to be closely related as well. Some cases of cognate paradigms will be presented that link these paradigms to each other.

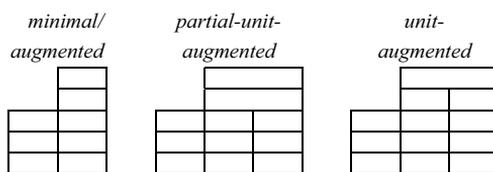


Figure 9.6: Major paradigmatic structures with an opposition between minimal and augmented inclusive

9.4.2 Cognate paradigms

Examples are presented from Nyulnyulan, Gunwingguan and Gé. The Nyulnyulan languages and the Gunwingguan languages are both from Australia. The examples in this section are rather biased towards cases from Australia, probably because the unit-augmented paradigm and the partial-unit-augmented paradigm are typical Australian paradigms, only rarely to be found elsewhere. The case of the Gé languages, spoken in South America, show that it is possible though unusual to find examples outside Australia.

Nyulnyulan (Australia)

The independent pronouns from the Nyulnyulan languages in Australia nicely illustrate the connection between the different variants of the minimal/augmented opposition. Shown below are the independent pronouns from the languages Nyulnyul (McGregor, 1996:23), Warrwa (McGregor, 1994:20-21) and Nyikina (McGregor, 1989:446, citing Stokes 1982). The Nyulnyul pronouns in (9.12) form a minimal/augmented paradigm; the Warrwa pronouns in (9.13) make up a partial-unit-augmented paradigm; and Nyikina in (9.14) has an unit-augmented paradigm. The morphemes are all clearly cognate. Only for the Nyikina morpheme *yarrjoo* is there no cognate in the other languages.

(9.12) NYULNYUL⁸

		yay	1+2
		yadir	1+2+3
1	ngay	yarrad	1+3
2	juy	kurr	2+3
3	kinyingk	(y)irr	3+3

⁸ Nyulnyul is a dying language. The independent pronouns as shown in (9.12) are dying with the rest of the language. Presently, the inclusive forms (1+2, 1+2+3) are hardly used by the last speakers of Nyulnyul. From older sources, the conclusion seems warranted that the paradigm as shown in (9.12) existed until a few decades ago. McGregor explains:

‘The 1&2 augmented forms [yadir] are absent from my corpus. The speakers who I worked with all employed the 1 augmented forms [yarrad] as general first non-singular pronouns, irrespective of whether or not the hearer was included. The full speaker very occasionally used the 1&2 minimal form [yay] for the speaker-hearer dyad. However, I was never able to elicit it systematically, and recorded it only a few times when it was uttered spontaneously.’ (McGregor, 1996:22)

(9.13) WARRWA

		<i>group</i>	<i>restricted group</i>	
		yawu		1+2
		yadirr		1+2+3
1	ngayu	yaarra	yaarra-wili	1+3
2	juwa	kurra	kurra-wili	2+3
3	kinya	yirra	jirra-wili	3+3

(9.14) NYIKINA

		<i>group</i>	<i>restricted group</i>	
		yayoo		1+2
		yarrjoo	yarrjoo-mirri	1+2+3
1	ngayoo	yarrga	yarrga-mirri	1+3
2	joowa	goorrga	goorrga-mirri	2+3
3	ginya	yirrga	yirrga-mirri	3+3

Gunwinguan (Australia)

Another example linking different variations on the minimal/augmented theme is found in the Gunwinguan family in Australia. The independent pronouns from Ngalakan (Merlan, 1983:71) and Rembarnga (McKay, 1978:28) are presented below. The Ngalakan pronouns form a partial-unit-augmented paradigm; the Rembarnga pronouns form a complete unit-augmented paradigm.

(9.15) NGALAKAN

		<i>group</i>	<i>restricted group</i>	
		yika?		1+2
		ṅurka?		1+2+3
1	ṅayka?	yirka?	yirka?-bira?	1+3
2	ṅiṅja?	ṅurka?	ṅurka?-bira?	2+3
3	niṅja?/jiṅja?	burka?	burka?-bira?	3+3

(9.16) REMBARRNGA

		<i>group</i>	<i>restricted group</i>	
		yũkkũ		1+2
		ngakorrũ	ngakorr-bbarrah	1+2+3
1	ngũũũ	yarrũ	yarr-bbarrah	1+3
2	kũ	nakorrũ	nakorr-bbarrah	2+3
3	nawũ/ngadũ	barrũ	barr-bbarrah	3+3

At first sight, these paradigms seem to be rather different. However, there are enough correspondences that indicate that the paradigms are cognate. The languages use different suffixes to mark these forms as independent pronouns (...-ka? in Ngalakan and ...-# in Rembarnga), but once these are removed, the unrestricted group pronouns are

rather similar, as shown in (9.17).⁹ However, for the present purpose, the crucial correspondence between Ngalakan and Rembarrnga is the use of the cognate suffixes ...-*biraʔ* and ...-*bbarrah* respectively. In Ngalakan, the use of this suffix leads to a partial-unit-augmented paradigm. In Rembarrnga, the addition of this suffix results in a complete unit-augmented paradigm.

(9.17)

	<i>Ngalakan</i>	<i>Rembarrnga</i>
1+2	yi-	yukk-
1+2+3	ɲur-	ngakorr-
1+3	yir-	yarr-
2+3	ɲur-	nakorr-
3+3	bur-	barr-

Gé (Brazil)

A final example showing a connection between the partial-unit-augmented paradigm and full unit-augmented paradigm is found among the prefixes of the Gé languages in Brazil. This is an interesting case, as it is the only example outside Australia that shows unit-augmented characteristics. However, it is not completely clear whether the non-singular marking is really grammaticalised. The sources are not explicit on this point. Most clearly, the Apinayé pronominal prefixes in (9.18) are described as a partial-unit-augmented structure (Callow, 1962:115). A cognate paradigm to Apinayé is found in the absolutive prefixes from Kayapo, shown in (9.19). This paradigm is described as a complete unit-augmented pattern, although the morphological status of morphemes in the paradigm is unclear. It might be the case that the number markers are morphologically independent.

(9.18) APINAYÉ¹⁰

	<i>group</i>	<i>restricted group</i>	
	paʔ-...		1+2
	mẽpaʔ-...		1+2+3
1	ič-...	mẽiçʔ-...	vaiç-... 1+3
2	a-...	mẽa-...	vara-... 2+3
3	iʔ-...	mẽʔ-...	vaʔ-... 3+3

⁹ The similarity between the roots of the pronouns in Rembarrnga and Ngalakan cannot be used to argue for a direct genetic relationship between the languages. In fact, the roots are much more widespread; they are clear instances of the proto-non-Pama-Nyungan pronouns reconstructed by Blake (1988:7), shown on page 213. The similarities as noted in (9.17) only argue for the inclusion of both languages within the non-Pama-Nyungan group of languages. The genetic relation between Rembarrnga and Ngalakan is established on other grounds.

¹⁰ Note that the names for the categories 'inclusive' and 'exclusive' are reversed in the grammar. Callow (1962:115:115, fn. 1) uses the name 'inclusive' for what is normally called 'exclusive' and 'exclusive' for what is normally called 'inclusive'.

(9.19) KAYAPO¹¹

		<i>group</i>	<i>restricted group</i>	
		(gu) ba-...		1+2
		(gu) mē ba-...	(gwaj) ba-...	1+2+3
1	i-...	mē i-...	ar i-...	1+3
2	a-...	mē a-...	ar a-...	2+3
3	∅-...	mē ∅-...	ar ∅-...	3+3

Presumably, the Kayapo paradigm is better interpreted as an only-inclusive paradigm with separate number marking. The non-singular markers *mē* and *ar* are written as independent prefixes, indicating that these markers are not a full part of the pronominal paradigm (Wiesemann, 1986b:368-369). The same prefix *me* is also found in Canela-Kraho, but in this case it is clearly described as an independent non-obligatory particle:

‘Number is sometimes expressed [in Canela-Kraho] by the particle *me* ‘plural’, usually where the referent is human and, more specifically, Indian.’ (Popjes & Popjes, 1986:185)

(9.20) CANELA-KRAHO

		pa-...	1+2
			1+2+3
1	i-...		1+3
2	a-...		2+3
3	ih-...		3+3

Another problem is the difference between the various forms of the inclusive. In Kayapo, the difference between the minimal inclusive (1+2) and the unit augmented inclusive (restricted 1+2+3) is marked by the independent pronouns *gu* and *gwaj*. Strictly speaking, this difference is not part of the prefixal paradigm, as it is marked by independent elements. It is only because *gwaj* is possibly a contraction of *gu* with the dual marker *ar* that I have included this opposition in this paradigmatic structure.¹² If the number marking is not part of the pronominal paradigm at all (like in Canela-Kraho), then the paradigmatic structure of Kayapo is of the same type as the one

¹¹ Actually, the dual forms from Kayapo are designated as ‘paucal’ by Wiesemann (1986b). She does not present any further argumentation why these forms should have paucal rather than dual reference. It is possible that she uses the name paucal in an idiosyncratic way to unify the referentially dual forms with the form (*gwaj*) *ba*..., which is strictly speaking a trial. Interpreted as such, the label ‘paucal’ by Wiesemann refers to the same category that I have called ‘restricted group’; the same group that is normally called ‘unit augmented’.

¹² The independent pronouns are ‘almost obligatory’ in case of the inclusive prefixes. These obligatorily used pronouns have been added between brackets.

‘In intransitive nominative clauses, the nominative [independent pronoun] is optional in the first, second and third persons, but almost obligatory (can be deleted in sloppy speech) in 1+2 person.’ (Wiesemann, 1986b:369)

shown in (9.20) for Canela-Kraho (Popjes & Popjes, 1986:175). More information on these languages is needed to decide on this issue.¹³

9.4.3 Summary

The three paradigms that were considered in this section are strongly connected. The minimal/augmented paradigm, the unit-augmented paradigm and the partial-unit-augmented paradigm are conceptually closely related (they all have an opposition between minimal and augmented inclusive). From the examples that were discussed in this section, it turns out that there are also close diachronic bonds between the paradigms. The various connections between the different variants of the minimal/augmented paradigm are summarised in Figure 9.7. They are added to the connections already established in the previous section. Note that the dual-inclusive/exclusive and the partial-unit-augmented are referentially identical paradigms.

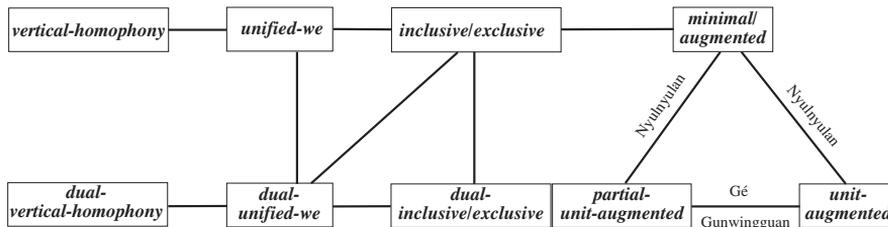


Figure 9.7: Second approximation of paradigmatic interconnectivity

Other direct connections probably do not exist. Only one theoretical link is left unaccounted for. The only missing link is between the unit-augmented paradigm and the inclusive/exclusive paradigm. This hypothetical link is shown in Figure 9.8.

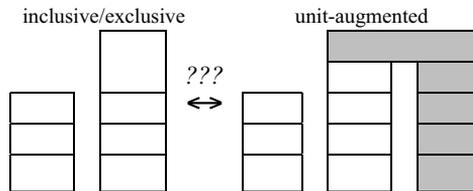


Figure 9.8: Not attested link between inclusive/exclusive and unit-augmented

I do not see how these paradigms could be directly connected. There are too many forms to be lost or added at once. This number of forms would not be a problem if it were not for the diversity of the referential values of these forms. For this transition,

¹³ A case that might shed some light on this situation is found in the Papuan language Amanab (Minch, 1991:31-32). In some crucial respects, the situation in the Amanab pronominal reference resembles the structure as described for the Gé languages. In Amanab, the normal form of the pronouns are 'proclitics' with an identical 4-way paradigm as found in the prefixes from Canela-Kraho in (9.20). The emphatic variants of the pronouns in Amanab add plural and dual marking to these 'proclitics'. Interestingly for the comparison with the Gé languages, the paradigmatic structure of these emphatic pronouns in Amanab is a dual-inclusive/exclusive paradigm, roughly comparable to the Apinayé paradigm in (9.18).

no unifying feature is suitable to account for all forms at once. It is far more probable that this transition is mediated by other paradigmatic structures. I expect that developments as sketched in Figure 9.8 will only happen in a few separate steps. However, this hypothesis is open to falsification.

9.5 Dual-3we as an intermediate

9.5.1 Preamble

In this section, the dual-3we pattern is connected to the other paradigmatic types. The dual-3we paradigms are not a homogenous group. At least five different paradigmatic structures were distinguished, which were only grouped together because of two shared characteristics. First, these paradigms have three different forms for ‘we’ (hence the name) and, second, these paradigms have a second and third person dual. Dual forms are also present amongst the three forms for ‘we’, but it is in this part of the paradigm that the variability of the dual-3we pattern prevails (see section 8.4). This group of similar but not completely identical paradigms turns out to be an intermediate structure, showing many different links to various other patterns. Because of this diachronic diversity it appears that the structural variability of the dual-3we paradigms is genuine and not an artefact of descriptive practices.

9.5.2 Cognate paradigms

Examples are presented from Yalandyic, Oregon, Daly, Burarran and Paman. Once again, these examples are biased towards languages from Australia. Only the case from a few languages from Oregon is not found among Australian languages. I expect this Australian bias to be an accidental skewing caused by the areal distribution of the dual-3we paradigms themselves. The majority of the examples of dual-3we paradigms are found in northwestern Australia and in southeastern New Guinea.

Yalandyic (Australia)

The first case connecting the dual-3we paradigm to the other paradigms is found in the Yalandyic family (part of Pama-Nyungan) in Australia. The independent pronouns in Kuku-Yalanji show a dual-3we paradigm (Oates & Oates, 1964:7). The neighbouring close relative Guguyimidjir has a dual-unified paradigm (de Zwaan, 1969:135). As can be seen from the paradigms in (9.21) and (9.22), almost all morphemes are identical. The main difference between the two paradigms is the presence (in Kuku-Yalanji) or absence (in Guguyimidjir) of a distinction between *ɲana* and *ɲanjin*. Only one of these, viz *ɲandan*, which is historically related to *ɲanjin*, is attested in Guguyimidjir as described by de Zwaan (1969). However, Haviland (1979) notes that there is a geographical difference between the use of either of the two forms in Guguyimidjir. The form *ɲandan* is used in the inland; resulting in the paradigm from de Zwaan (1969) as shown in (9.22). In the coastal dialect, the pronoun *ɲana* is used instead.¹⁴

¹⁴ Haviland (1979) also reports the existence of an exclusive dual *ngalliinh*, although this form is not in common use anymore.

'most people at the Hopevale Mission now use *nganhdhaan* in preference to the coastal form *ngana*'. (Haviland, 1979:65)

(9.21) KUKU-YALANJI

		<i>group</i>	<i>restricted group</i>		
		ŋana	ŋali(n)	1+2	
		ŋanjin		1+2+3	
1	ŋayu	yurra		yubal	1+3
2	yuudu	jana		bula	2+3
3	ñulu			3+3	

(9.22) GUGUYIMIDJIR

		<i>group</i>	<i>restricted group</i>	
		ŋandan	ŋali	1+2
		yura		1+2+3
1	ŋayu	yubal	yubal	1+3
2	nundu	dana	bula	2+3
3	nulu			3+3

Coos and Siuslaw (USA)

The two neighbouring languages Coos and Siuslaw in Oregon are probably, as far as is known today, not genetically related (Campbell, 1997: 119, 309-322). However, the pronominal inflections of these languages present a clear case of cognate paradigms linking the dual-3we pattern to the dual-inclusive/exclusive pattern. The Coos prefixes, presented in (9.23), have a dual-3we structure (Frachtenberg, 1922a: 321). The Siuslaw suffixes, presented in (9.24), form an dual-inclusive/exclusive paradigm (Frachtenberg, 1922b: 468). It is interesting to note that the cognate paradigms are prefixes in the one language, but suffixes in the other. Probably, both paradigms are historically derived from an independent set of pronominal elements that were prefigated in Coos and suffigated in Siuslaw (Siuslaw shows an overall preference for suffixes for some unknown reason). These 'proto' independent pronouns are lost in the mist of time.¹⁵

[continued from previous page]

'Most modern speakers do not make a distinction between inclusive ('you and I') and exclusive ('another person and I') in the first person dual, instead using *ngali* for an unspecified 1st person dual ('we two').' (Haviland, 1979:65)

¹⁵ The independent pronouns the lead to the affixal paradigms are *not* the extant pronouns of either language. The extant independent pronouns in both Coos and Siuslaw are derived from the affix paradigms. In Coos, the independent pronouns are made by adding the prefixes (9.23) to a root ...-*xka(n)* (Frachtenberg, 1922a: 395). In Siuslaw, the independent pronouns consist of the suffixes (9.24) added to the roots *ná*-... and *nix(ts)*-... (Frachtenberg, 1922b: 576).

(9.23) COOS

	<i>group</i>	<i>restricted group</i>	
1	n̥-...	ɬin-...	îs-... 1+2
2	e ^ɛ -...	cîn-...	xwîn-... 1+3
3	∅-...	îɬ-...	îc-... 2+3
			úx-... 3+3

(9.24) SIUSLAW

	<i>group</i>	<i>restricted group</i>	
1	...-n	...-nɬ	...-ns 1+2
2	...-nX	...-nxan	...-a ^u xûn 1+3
3	...-∅	...-tcî	...-ts 2+3
		...-nx	...-a ^u x 3+3

Daly (Australia)

For the next case I return to Australia. The Daly languages present a case for a link between the dual-3we pattern and the partial-unit-augmented pattern. The subject pronouns of Ngankikurungkurr are presented in (9.25). They form a dual-3we pattern with a gender distinction in the third person singular (Hoddinott & Kofod, 1988:94). The pronouns of the related Daly language Maranungku are shown in (9.26). These pronouns form a partial-unit-augmented pattern (Tryon, 1970:16). These two paradigms are not as clearly related as the above examples. However, on closer inspection, the differences are restricted to a few secondary aspects of the morphemes. First, the paradigms have different suffixes to mark number (...-gurr/rrim and ...-garri/riike in Ngankikurungkurr versus ...-tya and ...-tamata in Maranungku). When these differences are disregarded, the roots of the pronouns are almost identical. More important for the present discussion is the difference between the inclusive forms in the two paradigms. This difference makes these two paradigms a clear example of a cognate pair linking the dual-3we pattern to the partial-unit-augmented pattern.

(9.25) NGANKIKURUNGKURR

	<i>group</i>	<i>restricted group</i>	
1	ngayi	nayin 1+2	
2	nyinyi	ngagurr	ngagarri 1+3
3	nem/ngayim	nagurr	nagarri 2+3
		wirrim	wirrike 3+3

(9.26) MARANUNGKU

		<i>group</i>	<i>restricted group</i>	
		nganku		1+2
		kitya		1+2+3
1	ngany	ngatya	ngatamata	1+3
2	nina	nitya	nitamata	2+3
3	nankuny/ngankuny	witya	witamata	3+3

Burarran (Australia)

The next case links the dual-3we paradigm to the unit-augmented pattern. The examples presented are the independent pronouns from Burarra (Glasgow, 1984:15), shown in (9.27), and the independent pronouns from Ndjébanna (McKay, 1990:430), shown in (9.28). Only the masculine forms are shown here. The specialised feminine forms for the third singular and – quite unusual – for the dual forms are disregarded. The differences between the two paradigms are rather large. The similarity can best be recognised in the singular forms and in the forms for 1+2 and 3+3.

(9.27) BURARRA

		<i>group</i>	<i>restricted group</i>	
		ngarripa		1+2
		ngayburrpa	ngatippa	1+2+3
1	ngaypa	anagoyburrpa	anagotippa	1+3
2	nginyipa	birripa	bitippa	2+3
3	nipa			3+3

(9.28) NDJÉBANNA

		<i>group</i>	<i>restricted group</i>	
		ngárrabba		1+2
		ngúrrabba	ngirrikébbba	1+2+3
1	ngáyabba	njírrabba	njirrikébbba	1+3
2	njínjdabba	núrrabba	nirrikébbba	2+3
3	nakébbba	barrayabba	birrikébbba	3+3

Paman (Australia)

A final case of cognate paradigms, again from Australian languages, connects the dual-3we paradigm to the minimal/augmented paradigm. This time, the examples are from the Paman languages, a subgroup of Pama-Nyungan. A dual-3we paradigm is found in Wik-Munkan (Godfrey & Kerr, 1964:14). The conventional analysis of the three forms for ‘we’ is shown in (9.29a). However, in this case, a different lay-out seems more appropriate. This paradigm is shown in (9.29b), using an unconventional layout that suggests the possibility for a transition to the minimal/augmented paradigm.

(9.29A) WIK-MUNKAN

		<i>group</i>	<i>restricted group</i>	
		nampi	ngaali	1+2
		ngana		1+2+3
1	ngaya			1+3
2	ninta	niya	nipa	2+3
3	nila	tana	pula	2+3

(9.29B)

		<i>group</i>	<i>restricted group</i>	
		ngaali		1+2
		nampi		1+2+3
		ngana		1+3
1	ngaya			1+3
2	ninta	niya	nipa	2+3
3	nila	tana	pula	2+3

The related minimal/augmented paradigm is found in Uradhi (Crowley, 1983:352-356). The Uradhi pronouns, as shown in (9.30), may seem rather different, but they are closely related to the ones from Wik-Munkan. The difference is of a general phonotactic nature: Uradhi is one of the initial-consonant dropping languages of northern Australia. If the initial consonants from Wik-Munkan are removed, the Uradhi pronouns appear. Only the group pronouns *niya* (2+3) and *tana* (3+3) are lost. The former dual pronouns have taken over the complete non-singular marking in Uradhi.

(9.30) URADHI¹⁶

		ali(βa)	1+2
		ampu(la)	1+2+3
1	ayu(βa)	ana(βa)	1+3
2	antu(βa)	ipu(la)	2+3
3	ulu(βa)	ula(βa)	3+3

Finally, the minimal/augmented paradigm can be extended with dual marking, leading back almost full-circle to the paradigmatic starting point from Wik-Munkan. The newly added dual forms are found in Umpila, yet another Paman language (Dixon, 1980:355-356). As shown in (9.31), Umpila has a partial-unit-augmented pattern,

¹⁶ In the description of Uradhi (Crowley, 1983), the labels for the forms *ampu(la)* and *ana(βa)* are mixed up. They are glossed 'exclusive non-singular' and 'inclusive plural' respectively (Crowley, 1983:354-355). This is an error, as can clearly be seen from the example sentences 1-10 of text III (Crowley, 1983:397-398). In these sentences, the pronoun *ana-* is used in what is clearly an *exclusive* sense (excluding the addressee to whom the story is told). Terry Crowley adduces his 'sloppy proof-reading' (Crowley, personal communication) as an excuse. He clearly knew better, as the labels are correct in his description of the related language Anguthimri (Crowley, 1981:169-171).

based on the same forms as Uradhi in (9.30) – although in Umpila the initial consonants are preserved.¹⁷

(9.31) UMPILA

<i>group</i>		<i>restricted group</i>		
		ŋali		1+2
		ŋambula		1+2+3
1	ŋayu	ŋana	ŋana-baʔamu	1+3
2	ŋanu	ŋuʔula	ŋuʔula-baʔamu	2+3
3	nhulu	bula	bula-baʔamu	3+3

9.5.3 Summary

The dual-3we type turns out to be an intermediate case between various paradigms with a dual. The various connection that were presented in this section are summarised in Figure 9.9. I do not know of any direct connection between the dual-3we type and the inclusive/exclusive type. Perhaps some example will turn up in the future, but I do not expect this to happen.

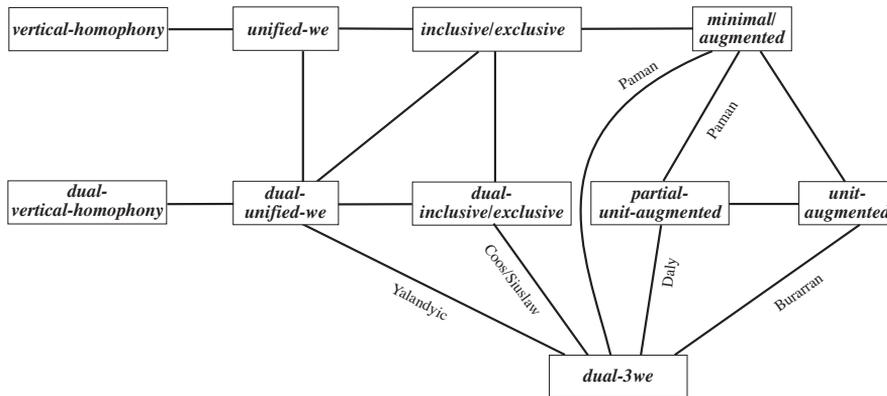


Figure 9.9: Complete version of paradigmatic interconnectivity

¹⁷ There is a strong affiliation between the pronouns from various Pama-Nyungan languages that were discussed in this chapter. All Pama-Nyungan examples are geographically located in the Cape York Peninsula, but show rather different paradigmatic structures. The pronouns from Kuku-Yalanji in (9.21) and Wik-Munkan in (9.29) are examples of the dual-3we type; Warrgamay in (9.3) and Guguyimidjir in (9.22) are examples of the dual-unified-we type; and Nyawaygi in (9.4) is an example of the dual-inclusive/exclusive type; Uradhi in (9.30) is an example of the minimal/augmented type; and finally, Umpila in (9.31) is an example of the partial-unit-augmented type. These paradigms are all clearly related to each other. The proto-Pama-Nyungan pronouns were probably of the dual-unified-we type (Dixon, 1980:Ch. 11; Blake, 1988:6). The development from there can probably be reconstructed as starting at the dual-unified-we, changing through the dual-3we pattern (with a possible side route over the dual-inclusive/exclusive), to the minimal/augmented pattern, and eventually ending in the partial-unit-augmented pattern. Dixon (1980:353-356) speculates about a reconstruction of a minimal/augmented paradigm for proto-Australian, partly on the basis of this paradigm from Umpila. The development as outlined here seems to be a much better proposal to explain the structure of the Umpila pronominal paradigm. There seems to be no need to reconstruct a minimal/augmented pattern for all Australian languages (Blake, 1988).

The overall impression is that the dual-3we pattern is clearly a dual pattern, mediating between other dual patterns. It is not related to the non-dual patterns such as the inclusive/exclusive or the unified-we. There is a link to the minimal/augmented pattern as this pattern is an intermediate structure between the paradigms with a dual and the paradigms without a dual. Another impression that arises from the many different connections of the dual-3we paradigms is one of inherent variability. Consequently, the large variety of dual-3we structures (cf section 8.4) is genuine and not an artefact of inaccurate descriptions.

9.6 Number marking incorporated

Two reductions of the complete set of paradigmatic connection will be made. As a result, the bewildering number of connections between the different paradigmatic types from Figure 9.9 will be reduced to a format that presents some promising typological patterns. This reduced paradigmatic interconnectivity will be brought together with the analyses from chapter 6. The result will be a cognitive map that shows the pathways through the similarity space of paradigmatic structure, directly linking paradigms that are cognitively near to each other.

The first cut to be made is that the dual-3we pattern is left out. This pattern is relatively common, but it is never a typical pattern for a whole group of genetically related languages. All examples are incidental cases within their close family. Moreover, the dual-3we pattern turns out to be a mediating structure between the other paradigmatic structure with a dual. The general impression is that the dual-3we pattern is an incidental ‘in between’ pattern, not a typical stable pattern for a human language (see section 9.5). Hence, it will be left out in the construction of a general outline of the dynamic possibilities of paradigmatic structures. Second, the partial-unit-augmented pattern will be combined with the dual-inclusive/exclusive pattern. Both patterns distinguish exactly the same referential categories. The main difference (and the reason to keep them apart in the first place) was the morphological structure. To obtain a general survey of the relations between the various paradigms, both patterns are collapsed into one type. The remaining set of paradigmatic connections is shown in Figure 9.10.

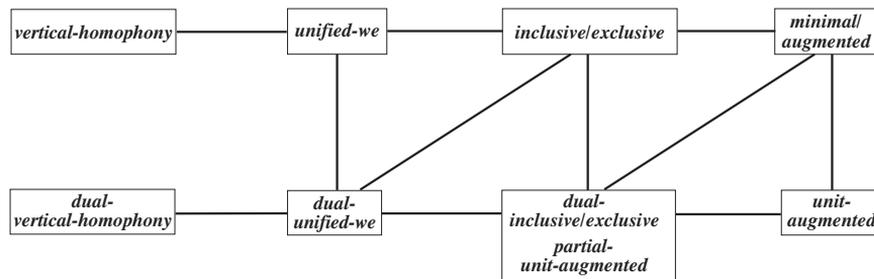


Figure 9.10: Reduced version of paradigmatic connectivity

This figure shows the connections that remain of the hypothesised connections between the Explicitness Hierarchy and the Dual Explicitness Hierarchy (cf Figure 9.2 on page 284).

Notably, the diagonals from top-left to bottom-right do not show up. These connections both involve multiple category-changes, and will consequently always be taken in at least two steps. First, the path between the unified-we paradigm and the dual-inclusive/exclusive paradigm needs both a addition of a dual and an addition of an inclusive/exclusive opposition. These will not be added to the paradigm in one process. More likely, the changes will follow a slight detour through other, less divergent, patterns. The other non-attested diagonal connects the inclusive/exclusive paradigm to the unit-augmented paradigm. This path needs both the addition of a dual and a separation of the two different inclusives. These changes will also not be conceived at once. In contrast, the two diagonals attested involve multiple category changes that can be accounted for by a single reanalysis. An inclusive can be reanalysed as a dual, trading, as it were, one category for the other. This reanalysis can only take place at the diagonals from upper right to lower left. The other diagonals can not be conceived as examples of reanalysis. The result of this remodelling of the paradigmatic interconnectivity is a cognitive map as shown in Figure 9.11. This map gives an impression of the similarity of the various paradigmatic structures. The structures that are connected through a line are cognitively similar. Diachronic change of the paradigmatic structure will tend to proceed along the lines of the cognitive map.

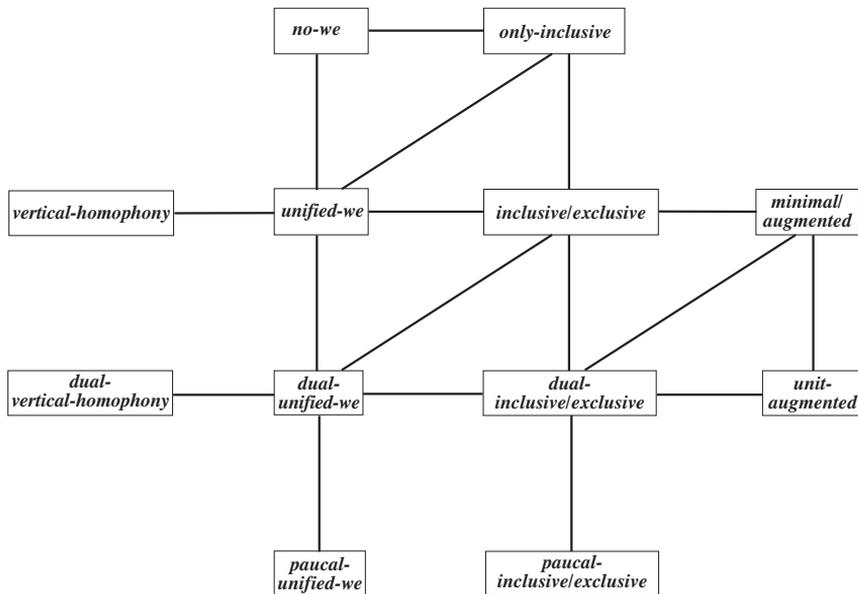


Figure 9.11: Cognitive map of paradigmatic interconnectivity

In this map, the connections from chapter 6 are added to the web of paradigmatic connection. The ‘no-we’ and the ‘only-inclusive’ pattern are shown on top of Figure 9.11. These connections fit in perfectly with the others. In the top part of the figure,

the same diagonal is found as in the dual part, linking two paradigmatic types through reanalysis. Both the only-inclusive and the unified-we paradigm have only one specialised element for ‘we’. This element has different referential value in both paradigms, but these two are diachronically related. The other diagonal, hypothetically linking no-we to inclusive/exclusive, would need two forms to arise or be lost in one change. This connection is not attested. Probably, a more gradual path is taken, intermediated by one of the other types.

On the lower side of the figure, the two main paradigms with paucal marking are added. An example of cognate paradigms linking the dual-unified-we to the paucal-unified-we is found in the Lower Sepik family from New Guinea. The languages Murik, Yimas and Chambri from this family have a paucal-unified-we paradigm, but the language Karawari has only a dual-unified-we paradigm. Foley (1986:219-221) argues that the proto-Lower Sepik pronominal paradigm had a paucal-unified-we paradigm. An example of cognate paradigms linking the dual-inclusive/exclusive to the paucal-inclusive/exclusive can be found in the Malaita family (a sub group from the Oceanic languages) from the Solomon Islands. The southern languages Kwaio, Sa’a, Langalanga and Lau have a paucal-inclusive/exclusive paradigm, but the northern language To’abaita has a dual-inclusive/exclusive paradigm. Simons (1986:33-34) argues that the proto-Malaita pronominal paradigm had a paucal-inclusive/exclusive paradigm.¹⁸

9.7 Hierarchies

The labels in the cognitive map only refer to the structure of the first person complex. These labels describe the kind of marking in the paradigm that is used for the same reference as the English pronouns ‘we’. This is an abstraction from the rather large variability of paradigmatic structure. Various kinds of homophony were attested in rather large quantities, but this variation did not find its way into the cognitive map. Only the twelve different structures of the first person complex remained as salient structures to be included in the cognitive map. Prototypical examples for each of the twelve types of the first person complex are shown in Figure 9.12 to illustrate the paradigmatic structure of each case. The central four paradigms in the cognitive map are the structures that were coined as ‘the four systems that are more frequent than the others’ by Ingram (1978:219). These patterns are indeed the most frequent ones, but the other paradigmatic structures are also commonly found and should not be put aside too easily.

¹⁸ These paradigms with a paucal are not really common: the only examples of such paradigms are found in ‘Melanesia’: New Guinea and the surrounding smaller islands. The paradigms with a paucal have a comparable status to the dual-vertical-homophony paradigms and the unit-augmented paradigms, which are also only found in a restricted set of cases. These two paradigmatic structures are attested in southeast New Guinea and in northwest Australia respectively. In general, the four lower outmost paradigmatic structures in Figure 9.11 are all found in a strongly restricted region: southeastern New Guinea, northwestern Australia and the nearby Oceanic languages.

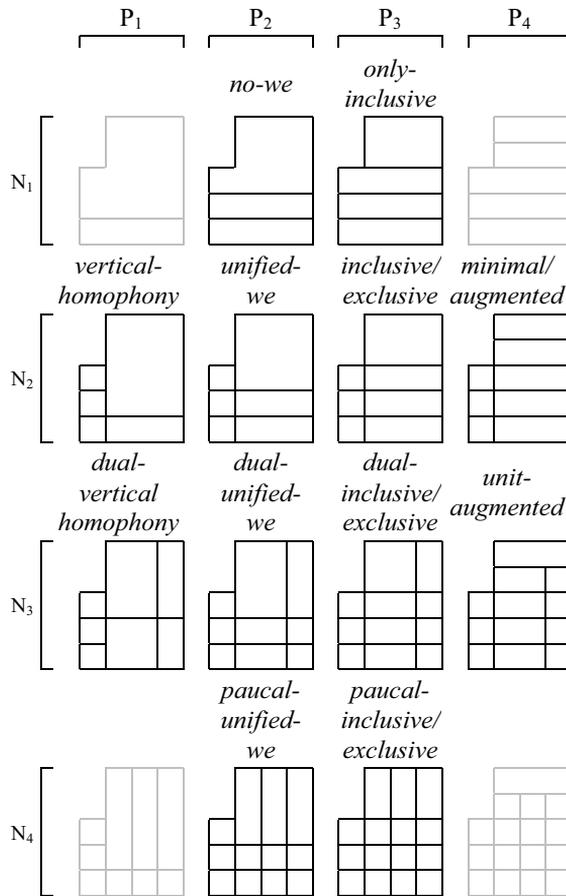


Figure 9.12: Examples of the paradigmatic structures in the cognitive map

The vertical dimension of Figure 9.12 can be described as a hierarchical tree of number oppositions. The four rows will be discussed in turn from top to bottom. The paradigms that do not distinguish number are located in the upper row (stage N₁). The inclusive that is grammaticalised in the only-inclusive paradigm is not a special kind of number marking (an impression that might arise from the old name ‘first person plural inclusive’), but a kind of person marking. The second row (stage N₂) is characterised by a consistently marked opposition between singular and group marking. In fact, there are many paradigms that fall in between stage N₁ and N₂. The opposition between singular and group is not marked consistently throughout the paradigm if some horizontal homophony is present. All paradigms that are only partly horizontal homophonous fall in between stage N₁ and N₂. The third row (stage N₃) consists of the paradigms with ‘restricted group’ marking. Most of these paradigms have dual forms, but the unit-augmented paradigm (on the far right of the third row) has a trial inclusive that behaves parallel to the dual forms in the other paradigms. The label ‘minimally restricted group’ has been proposed to cover both dual and inclusive trial of referential values. The fourth and final row (stage N₄) consists of the paradigms with paucal marking. This hierarchical tree of oppositions is shown in Figure 9.13.

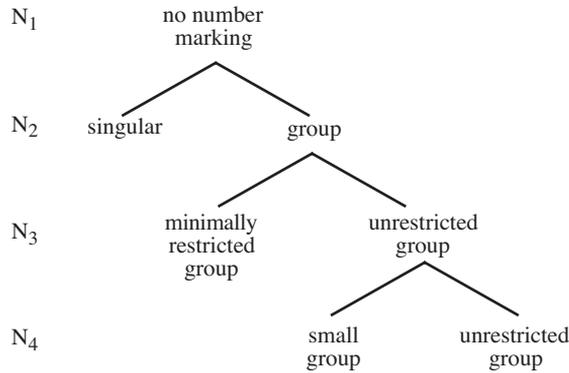


Figure 9.13: Hierarchical tree of number oppositions

The horizontal dimension of Figure 9.12 can be described as a hierarchical tree of person oppositions. The four columns will be discussed in turn from left to right. The first column (stage P₁) consists of paradigms that have intermingled person categories. The two paradigmatic structures presented in the first column of Figure 9.12 are the most common kinds of incomplete person marking. However, many more patterns belong to this column, for example paradigms with a first + third person vertical homophony and paradigms with singular homophony. The paradigms that are presented in the second column (stage P₂) make a three-way opposition between ‘speaker and addressee excluded’ versus ‘speaker excluded and addressee included’ versus ‘speaker included’ throughout the paradigm. These three categories are traditionally called ‘third person’, ‘second person’ and ‘first person’, respectively. The third column (stage P₃) is characterised by an opposition between ‘addressee excluded’ and ‘addressee included’ within the category ‘speaker included’. The resulting forms are traditionally known as ‘first person plural exclusive’ and ‘first person plural inclusive’, respectively. The final column (stage P₄) specifies a minimally restricted inclusive and an unrestricted inclusive. This opposition is also known as ‘minimal inclusive’ versus ‘augmented inclusive’. The resulting hierarchical tree of oppositions is summarised in Figure 9.14.

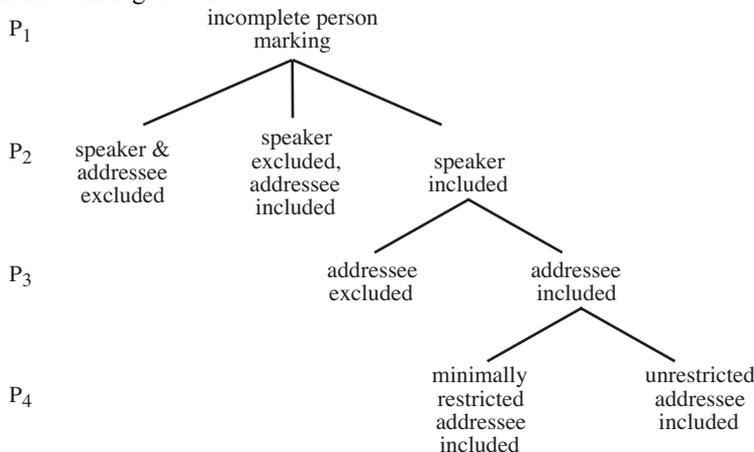


Figure 9.14: Hierarchical tree of person oppositions

The person and number dimensions are not independent of each other. Specifically, number opposition N_3 and person opposition P_4 describe exactly the same opposition: ‘minimally restricted’ versus ‘unrestricted’. This correspondence between the two hierarchies explains the low variability of the minimal/augmented paradigms. The paradigms that have an opposition between a minimal inclusive (1+2) and an augmented inclusive (1+2+3) hardly show any structural variation. There are almost no cases with vertical homophony or horizontal homophony among these paradigms. This observation resulted earlier in a description of the opposition between minimal and augmented inclusive as the ‘icing on the cake’ for a pronominal paradigm. Why this is so can now be understood. The opposition between ‘minimally restricted’ and ‘unrestricted’ is in fact an opposition that refers to the cardinality of the group. It only makes sense to apply this opposition when there is a difference between singular and non-singular. The opposition between ‘minimally restricted inclusive’ and ‘unrestricted inclusive’ (stage P_4) can only be marked in a paradigms when there is a completed opposition between singular and group (stage N_2). The combination of stage P_4 with stage N_1 does not make sense. I do not expect the hypothetical structure as shown in the upper right corner of Figure 9.12 to show up after more research. The other hypothetical structures in the corners in Figure 9.12 could very well exist. However, they are highly marked structures on the fringes of the possibilities of human language structure.

9.8 Conclusion

In this chapter, a collection of cognate paradigms has been presented to test the hypothesis that pronominal paradigms change through time and space along the lines of the previously established typological hierarchies. In previous chapters, the structure of pronominal paradigms has been shown to be ordered along the lines of two hierarchies: the Explicitness Hierarchy and the Dual Explicitness Hierarchy. These hierarchies were taken in this chapter as a hypothesis for the possible changes of paradigmatic structure. The hypothesised connections are repeated here in Figure 9.15.

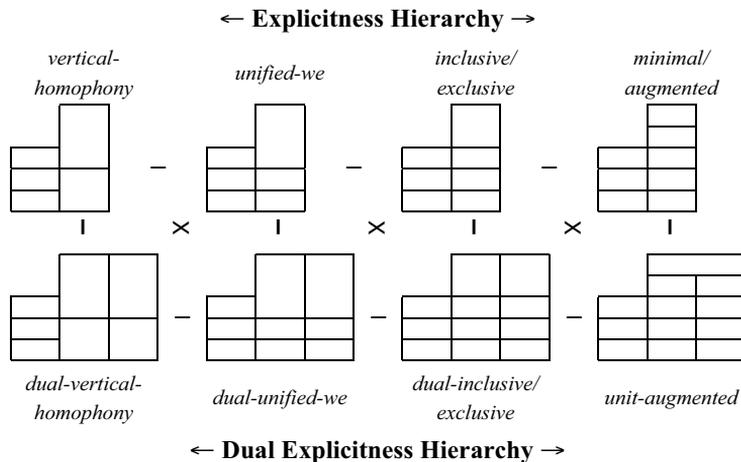


Figure 9.15: Hypothesised connections along the Explicitness Hierarchies

To test this hypothesis, examples of cognate paradigms have been collected. Cognate paradigms are pronominal paradigms from closely related languages that are clearly related morphophonologically, but that differ in their paradigmatic structure. Numerous cases have been presented, which show that many of the neighbouring paradigms on the hierarchies are closely related. The hypothesised connections are almost all attested. Only a few connections are missing from the far left side of Figure 9.15. On the basis of the connections attested as discussed in this chapter combined with the results from chapter 6, a cognitive map has been constructed that shows how the various major paradigmatic structures are related to each other. This cognitive map is shown here in Figure 9.16, ordered to the number of forms in the paradigm that are to be translated into English as ‘we’.

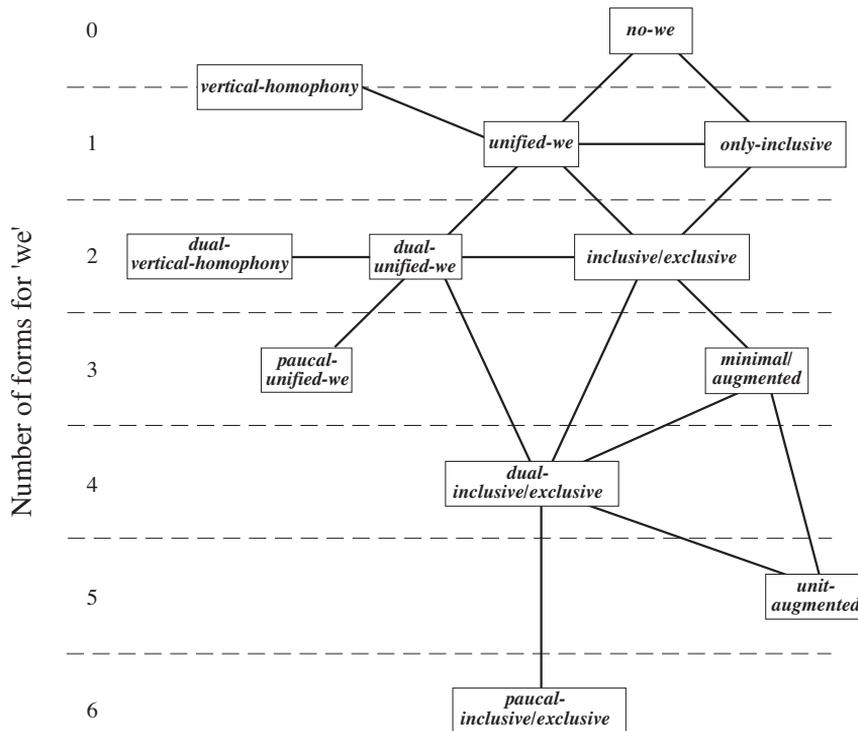


Figure 9.16: Cognitive map ordered to the number of forms for ‘we’

This cognitive map summarises a few of the major results of the present work. First, one of the findings is that the structure of the first person complex is a strong predictor for the variability of the rest of the paradigm (see especially section 5.7). Also, not all possible structures of the first person complex are attested equally frequently; there are clear preferences for specific markedness structures of the first person complex (see especially section 3.6). In the cognitive map, the structure of the major types of the first person complex are shown. This survey summarises the major possibilities of person marking paradigms, as characterised by a central feature of such paradigms: the first person complex. Second, the various types in the cognitive map are interconnected. These connections represent the similarity between the paradigmatic struc-

tures. Structures that are similar are connected by a direct line. The similarity has been established by a combination of typological and diachronic reasoning in this chapter and in chapter 6. The similarity is basically typological, that is, the lines as drawn in the cognitive map show an accurate representation of the variation attested among the world's languages. However, this similarity-space can very well be interpreted as the paths of 'least resistance' for diachronic change. It is surely not necessary for change to follow along the lines as shown in the cognitive map, yet it is likely. The connections between the various paradigmatic structures as shown in the map are the expected kinds of change. If another change happens, this is in need of additional explanation.

Chapter 10

Finale

Summary and prospects

10.1 Summary

10.1.1 Preliminaries

The structural variation among the world's languages is overwhelming. Yet, this variability should not be considered a nuisance to linguistic inquiry, but a potential source of insight into the possibilities of human language. Looking for a balance between gross generalisations and extreme relativism, I have attempted, in the present investigation, to catalogue the existing variability of the paradigmatic structure of person marking. Certain structures turned out to be more common than others. However, the existing variation shows a continuum from commonly occurring paradigmatic structures, through a set of ever less common structures up to structures that are not attested among the world's languages at all. Nothing seems to be impossible, although certain structures are clearly less probable.

In this study, the reference to speech-act participants has been the subject of cross-linguistic investigation. Such reference is performed by words like the English *I* and *you*, but also by inflectional elements like the English suffix ...-s, which marks for 'not speaker nor addressee'. More specifically, only the reference by specialised shifters has been included in this study. Words like *I* or *you* are such specialised linguistic elements with shifting reference depending on who speaks them. *I* always refers to the person who is presently speaking; *you* always refers to the addressee. When the turn in a conversation is taken by another participant, the reference of these words changes. It is impossible for such elements not to shift in reference. This definitional property is important to delimit the cross-linguistic investigation as reference to speech-act participants can be performed in many ways. A father can refer to himself by *daddy* or a teacher can refer to a student by using the proper name, like in *How would John explain this?* The usage of regular nouns ('daddy') or proper names ('John') for reference to speaker or addressee is a stylistic device that is rather restricted in English, but in many other languages it is much more widely spread. Such linguistic devices that can be used for speaker or addressee reference, but that do not necessarily perform that function, are excluded from the present investigation. Specialised shifters that are used for speech-act reference are called PERSON MARKERS (see sections 1.2.2 and 1.2.4).

Person markers are bound into a paradigm. A paradigm of person markers consists of a set of elements that are syntagmatically equivalent in the structure of a language. The elements in a paradigm take, so to speak, the same place in the overall construct that is called a language. They have to belong together both functionally and structurally. Functionally, the elements have to be, for example, all subject markers, or dative markers, or whatever markers, as long as all elements perform the same function. Structurally, the elements have to be either all independent or all inflectional; and if they are inflectional, then they have to be all prefixes or all suffixes, filling the same inflectional ‘slot’ on a predicate. Each paradigm of person markers should at least have an opposition between (singular) speaker and addressee reference, as, for example, the English subject pronouns *I* and *you*. All person markers that belong to the same paradigm form the PARADIGMATIC STRUCTURE; for the English subject pronouns this includes the elements *I*, *you*, *he*, *she*, *it*, *we* and *they*. Note that gender and honorifics, among other pronominal dimensions, are disregarded in the present investigation (see sections 1.2.3 and 1.2.5).

10.1.2 Person and number

The paradigmatic structure of person marking is traditionally analysed in two orthogonal dimensions: person and number. In this tradition, the dimension ‘person’ consists of three persons – first, second and third – and a possible inclusive/exclusive opposition in the first person plural. The dimension ‘number’ consists of the categories singular, plural, dual, trial, paucal, and possibly a quadral. I have argued that this analysis is not ideal as an approach to the cross-linguistic variation in person marking. Semantically, the notion ‘plural’ is not suitable for words like ‘we’ because ‘we’ is not the plural of ‘I’. Morphologically, the ‘plural’ person markers are, in the far majority of cases, not derived from the singular markers. Moreover, the inclusive/exclusive opposition is only relevant in the first person plural, which indicates that ‘person’ and ‘number’ and not really orthogonal dimensions (see section 3.3).

I have proposed a different analysis of person marking. The categories that are traditionally called ‘plural’ have been reanalysed as ‘groups’ of singular participants. These groups are inherently plural as they consist of more than one participant. Yet, it is the KIND of participants in a group that is important, not the NUMBER of participants. Five different kinds of groups turned out to be cross-linguistically viable categories. These five ‘plural’ groups and the three singular categories together, as shown in (10.1) below, form the basic grid for the typology of the paradigmatic structure of person marking (see sections 3.4 and 3.5).

(10.1)

- 1: speaker (‘I’)
- 2: addressee (singular ‘you’)
- 3: neither speaker nor addressee (‘he’, ‘she’, ‘it’)
- 1+2: speaker and addressee (minimal inclusive ‘we’)
- 1+2+3: speaker and addressee and others (augmented inclusive ‘we’)
- 1+3: speaker and others (exclusive ‘we’)
- 2+3: addressee and others (plural ‘you’)
- 3+3: multiple others (‘they’)

Graphically, these eight categories are depicted as a two-dimensional paradigm as shown in Figure 10.1. This graphical representation is used throughout the present study to allow for an easy comparison between the various paradigmatic structures.

		<i>'non-singular'</i>			
		1+2	<i>minimal inclusive</i>	}	<i>inclusive</i>
		1+2+3	<i>augmented inclusive</i>		
<i>speaker</i>	1	1+3	<i>exclusive</i>	}	<i>first person complex</i>
<i>addressee</i>	2	2+3	<i>second person plural</i>		
<i>other</i>	3	3+3	<i>third person plural</i>		

Figure 10.1: The paradigmatic grid that has been used for the typological classification

As a result of this recategorisation from 'plural' into 'group', the singular and the group categories are considered now as 'unmarked for number'. The dual, trial, and the other traditional number categories, remain as those categories that are 'marked for number'. There are a few cases that show signs of a markedness reversal; in these cases, the dual is unmarked relative to the group marking. However, such reversals are only found in a small set of exceptional cases (see section 7.3). Among the number categories, the dual is clearly most prolific cross-linguistically. Other categories, like trial, paucal or quadral, occur only sparingly and in a restricted area (roughly the area that is known under the name of 'Melanesia'). Moreover, it is questionable whether the trial really exists as a grammaticalised person category. I have argued that, on the basis of the published accounts, it seems better to interpret all trials as being in fact paucals (see section 7.4).

10.1.3 Paradigmatic structure

The typological variation in the markedness structure of these categories is large. In total, 98 different paradigmatic structures have been described in this investigation; if even more languages had been included, this number would surely rise.¹ This large variety of attested structures already indicates that the often assumed basic '6-way' paradigm (three persons in two numbers) is not the only possibility. The variation attested was discussed in a few separate parts throughout the investigation. First, the markedness structure of the three singular categories (speaker, addressee and other) was investigated in sections 2.3 and 2.4. All possible kinds of homophony between the three categories are attested. Second, the markedness structure of the first person complex (the group categories that include at least the first person) was investigated in sections 3.5 and 3.6. Here, the first restrictions on the typological possibilities of human language are observed. Only 10 out of 15 possible structures are attested, and 5 out of these 10 are found in only one or two instances. The five remaining 'common' structure of the first person complex are shown in (10.2).

¹ These 98 different paradigmatic structures are described at different places in the discussion. The first 6 cases are discussed in section 3.6.6; then 55 cases are discussed in chapter 4; the next 5 cases are discussed in section 7.4; finally, 32 cases are discussed in chapter 8. A list of the various structures is presented in the Appendices B and C.

- (10.2)
- | | |
|-----------------------------|--|
| <i>no-we:</i> | there is no specialised form for ‘we’ in the paradigm. The same morpheme that is also used for speaker reference is used for the first person complex |
| <i>only-inclusive:</i> | there is one morpheme in the paradigm that translates as ‘we’, but only in the inclusive sense. The exclusive ‘we’ is marked by the same morpheme that is also used for reference to the speaker |
| <i>unified-we:</i> | there is one morpheme in the paradigm that encompasses all reference of the first person complex, like the English pronoun <i>we</i> |
| <i>inclusive/exclusive:</i> | there are two morphemes for ‘we’, one for inclusive reference and one for exclusive reference |
| <i>minimal/augmented:</i> | there are three morphemes for ‘we’, one for minimal inclusive ‘we’ (only speaker and addressee), one for augmented inclusive ‘we’ (speaker addressee and other) and one for exclusive ‘we’. |

On the basis of this classification in five major types, a complete discussion of the paradigmatic variability attested was presented in chapter 4. A summary of this variation was given in section 4.8. There turns out to be a continuum, ranging from commonly attested structures, over semi-common and rare structures, to non-attested structures. The attested frequencies do not show a clear cut-off point between typologically ‘basic’ and ‘abnormal’ paradigmatic structures. Everything seems to be possible, yet some structures are clearly more probable than others.

The marking of specialised number categories is discussed in chapter 7. The only really common number category in the domain of person marking is the dual. Roughly summarised, only the last three types from (10.2) occur with dual morphemes. A complete discussion of the paradigmatic variation of the dual is presented in chapter 8. A summary of this variation is given in section 8.9. Again, a continuum of frequency is attested: some paradigms are clearly more common than others, but there is no clear cut-off point between ‘basic’ and ‘abnormal’ paradigmatic structures. In general, the most common paradigmatic structures with a dual are equivalent to the most common structures without a dual.

10.1.4 Hierarchies

The question remains why certain paradigmatic structures are attested more commonly than others. To find a rationale behind this distribution, the internal structure of the paradigms is scrutinised. The paradigms without specialised number marking are analysed in chapter 5. The same analyses are repeated for the paradigms with dual marking in section 8.8. The paradigms with a dual behave roughly equivalent to the paradigms without a dual. The generalisation from both chapters are combined and summarised in sections 9.6 and 9.7. Here, I will quickly summarise the analyses from chapter 5. The result of these analyses is a classification of important characteristics of paradigmatic structure. Two sets of interdependent paradigmatic parameters can be distinguished. These two sets can be summarised as hierarchies that describe most of

the variation of the internal structure of the paradigm: the Explicitness Hierarchy and the Horizontal Homophony Hierarchy.

The Explicitness Hierarchy describes the order in which a particular set of oppositions in a paradigm is grammaticalised (see section 5.5). The more of these oppositions are grammaticalised, the more explicit is the reference to the various referential person categories. At the highest level, all eight person categories, as shown in (10.1) above, are differentiated by separate morphemes. On the way to a less explicit paradigmatic structure, there is a strict order in which the various categories are combined into the reference of one morpheme. The various stages on this hierarchy are described in (10.3) below. The first oppositions to give way are the oppositions in the first person complex. Minimal inclusive (1+2) and augmented inclusive (1+2+3) are combined first, and next this cluster is combined with the exclusive (1+3) to yield a unified ‘we’ (like the English pronoun *we*). Only after the first person complex has been united, can other non-singular categories be combined into clusters of referential categories (this is called ‘vertical homophony’). Finally, only if the non-singular reference is thus strongly reduced in its referential explicitness, can some of the singular categories also be combined into the marking of one morpheme. Such a singular homophony (like, for example, the zero marking in the English inflection, which marks for singular speaker and addressee) is the lowest rung of explicitness for a paradigm of person.² Sometimes it is even questionable whether it is really ‘person’ that is marked on this lowest rung.

(10.3) Explicitness Hierarchy

- | | |
|--|---|
| <i>most explicit:</i> | <ul style="list-style-type: none"> – all eight person categories, as shown in (10.1) above, are differentiated – minimal inclusive and augmented inclusive are combined in the marking of one morpheme – all categories of the first person complex are combined into the marking of one morpheme – the second or third person plural is marked together with the ‘unified’ first person complex (‘vertical homophony’) |
|  | <ul style="list-style-type: none"> – on top of that, some of the singular categories are combined into the marking of one morpheme (‘singular homophony’) |
| <i>least explicit:</i> | |

The Explicitness Hierarchy is graphically represented by a few selected paradigmatic structures in Figure 10.2. The paradigmatic structures that are shown here follow the grid as presented in Figure 10.1 above.

² The Explicitness Hierarchy describes a cross-linguistically salient conceptualisation of the notion ‘richness’ of a paradigm. The notion ‘richness’ has been invoked in the generative literature to explain why certain languages allow *pro*-drop and others do not. Languages that have a ‘rich’ inflectional subject paradigm, so the reasoning goes, do not need to add the independent pronoun. However, this reasoning does not work with the Explicitness Hierarchy. Even paradigms with a singular homophony, which are on the lowest rung of the Explicitness Hierarchy, are still found to allow *pro*-drop in many of the world’s languages (see section 2.5).

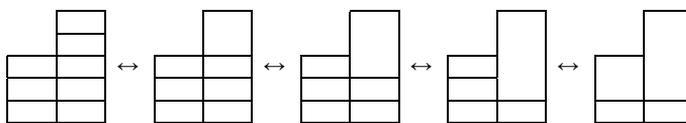


Figure 10.2: Selected examples from the Explicitness Hierarchy

The other main generalisation is the Horizontal Homophony Hierarchy (see section 5.6). Horizontal homophony is characterised by a homophony between two categories, one of which is singular and one of which is non-singular (ie a ‘group’ category). With three singular and five group categories, there are 15 theoretically possible kinds of horizontal homophony. However, four kinds of horizontal homophony account for the far majority of cases: the third person singular is homophonous with the third person non-singular ($3 \rightarrow 3+3$); the second person singular is homophonous with the second person non-singular ($2 \rightarrow 2+3$); and the first person singular is homophonous with the exclusive and/or the inclusive first person non-singular ($1 \rightarrow 1+3$ and $1 \rightarrow 1+2/1+2+3$). Other kinds (called ‘diagonal homophony’) are only attested in a few exceptional examples. Moreover, the four common kinds of homophony are connected by a Horizontal Homophony Hierarchy, as shown in (10.4).

(10.4) Horizontal Homophony Hierarchy

No Homophony < ($3 \rightarrow 3+3$) < ($2 \rightarrow 2+3$) < ($1 \rightarrow 1+3$) < ($1 \rightarrow 1+2/1+2+3$).

This Hierarchy is graphically represented by a few selected paradigmatic structures in Figure 10.3. The paradigmatic structures that are shown here follow the grid as presented in Figure 10.1 above.

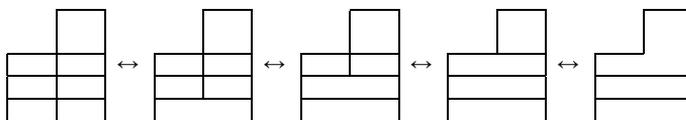


Figure 10.3: Selected examples from the Horizontal Homophony Hierarchy

These hierarchies were established on the basis of an investigation of inflectionally and independently marked paradigms. Beforehand, no separation was made between different forms of the morphological marking. Simply all kinds of paradigms of person were included. As a result of this broad perspective, it has been shown that both the Horizontal Homophony Hierarchy and the Explicitness Hierarchy correlate with the morphological status of the paradigms. The higher on both hierarchies, the larger the proportion of morphologically independent paradigms. Paradigms that are low on either of the two hierarchies are predominantly inflectionally marked. It should not be concluded too easily that this correlation means that large person paradigms are therefore independently marked. This correlation describes a cross-linguistic generalisation, and individual languages can, and often will, behave rather differently (see section 5.8).

10.1.5 The concept of Pure Person

A central cut-off point for the paradigmatic structure of person marking is the inclusive/exclusive opposition. The paradigms without an inclusive/exclusive opposition

allow for much greater liberty in combining different referential categories into the marking of one morpheme. Once there is an inclusive/exclusive opposition in the paradigm, the possible variation is strongly constrained. This difference is graphically represented in Figure 10.4 (taken from section 5.7). Paradigms without an inclusive/exclusive opposition are represented on the left; paradigms with an inclusive/exclusive opposition are represented on the right (the opposition between a minimal and an augmented inclusive is disregarded). In the pictures, the various person categories are connected by lines that represent the number of paradigms in which the two categories are found to be homophonous. The thicker a line, the more examples of this kind of homophony are attested in the sample. If there is no line between two categories, this means that the categories are never homophonous (or attested in maximally one case).

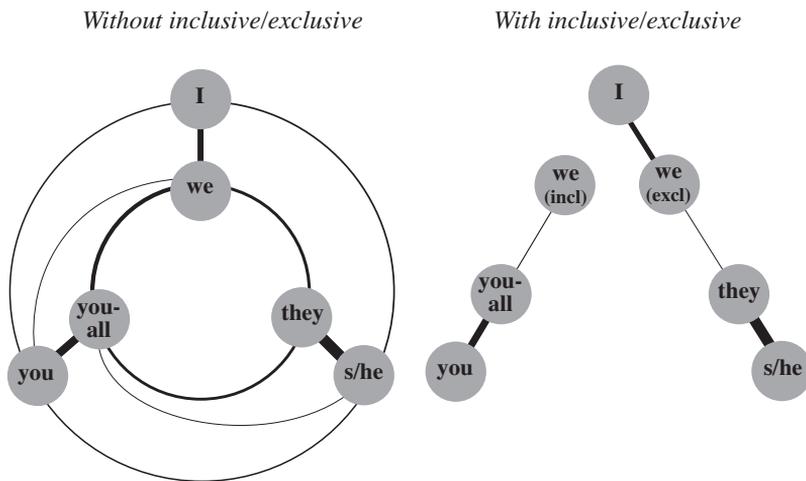


Figure 10.4: The impact of an inclusive/exclusive opposition for the paradigmatic structure

In general, the two pictures show a rather different geometry. Yet, there is hardly a difference as to the occurrence of horizontal homophony. In both pictures, ‘I’ is connected to ‘we’ by relatively thin lines; likewise, ‘you-singular’ is connected to ‘you-plural’ and ‘he/she/it’ to ‘they’.³ The major difference between the overall geometry of the two pictures is found in the other kinds of homophony. In the left hand side picture, there are 12 theoretically possible kinds of homophony left when horizontal homophony is disregarded.⁴ Of these 12 theoretical possibilities, 8 kinds of homophony are actually attested. In contrast, in the right hand side picture, there are 16 theoretically possible kinds of homophony left.⁵ However, in this case only two kinds of

³ A clear difference is the kind of horizontal homophony in the first person: in case of an inclusive/exclusive opposition, the first person singular is possibly homophonous with the exclusive, not with the inclusive.

⁴ With six categories, there are $5+4+3+2+1=15$ possible connections. The three kinds of horizontal homophony are subtracted, yielding 12 possibilities.

⁵ With seven categories, there are $6+5+4+3+2+1=21$ possible connections. The four kinds of horizontal homophony (one of which is not attested) are subtracted, yielding 17 possibilities. Also, the theoretical

homophony are actually attested. Moreover, these two connections in the right picture are only attested in a few cases (as can be seen from the rather thin lines). To summarise, the paradigms with an inclusive/exclusive opposition show much less occurrence of homophony between person categories. They predominantly allow for horizontal homophony, which is constrained by the Horizontal Homophony Hierarchy.

These generalisations can all be subsumed under a theoretical concept that I call ‘Pure Person’. Paradigms of person can be more or less ‘pure’ in their conceptualisation of the dimension ‘person’. The marking of ‘person’ is directed to the role in the speech-act, most importantly to the roles ‘speaker’ and ‘addressee’. By definition, the paradigms that were included in this investigation mark at least some person oppositions. However, the centrality of the person categories to the paradigm varies. Within the marking of person, the various roles are valued as more or less central to the speech-act. Cross-linguistically, there is a strong person hierarchy placing speaker above addressee above others. In some languages in some specific instances, this hierarchy might be ordered differently, but generalising over the attested diversity among the world’s languages, the hierarchy is strongly substantiated by the Horizontal Homophony Hierarchy. Moreover, the person hierarchy has been found to regulate many more clines in the diversity of structures (see sections 5.3.3 and 5.5.3). Also, the Explicitness Hierarchy describes a continuum between greater or lesser importance of the person categories. The higher on the Explicitness Hierarchy, the purer the marking of person. The difference between inclusive and exclusive is a major breaking point on this continuum. This opposition is the crown on the marking of person, as the difference between the combination speaker/addressee (inclusive) and speaker/others (exclusive) is marked overtly in the paradigm. Only in cases of total absence of horizontal homophony, is it possible for a paradigm to add an even more stringent opposition between speaker/addressee/no-other (minimal inclusive) and speaker/addressee/other (augmented inclusive). Finally, it has been shown that the purer the marking of person in a paradigm, the larger the proportion of independently marked paradigms (see section 5.8). The purer a paradigm of person, the more especially this paradigm is devoted to the marking of person. A Pure Person paradigm becomes a specialised sub-system of a language, functioning to some extent independently of the rest of the language.

10.1.6 Diachronic interpretation

Typological hierarchies might be connected to the possible directions of language change. The hierarchies from the present investigation could, for example, be the result of certain restrictions on possible changes. To test whether the typological hierarchies may serve as hypotheses for language change, examples of cognate paradigms are collected in chapter 6 and 9. The idea behind such cognate paradigms is that changes in paradigmatic structure may perhaps be visible in the small differences between closely related languages. Examples of such cognate paradigms present a window on the dynamics of language change. By compiling a large collection of ex-

[continued from previous page]

homophony between the inclusive and the exclusive is not counted as a possibility, because this opposition is the defining characteristic of this picture.

amples of cognate paradigms, an impression can be obtained of which paradigmatic structures are closely related (for methodological considerations, see sections 6.2 and 9.2). It turned out that the Horizontal Homophony Hierarchy does not present a good hypothesis of diachronic change. Closely related paradigms appear to jump up and down this hierarchy with great strides (see section 6.3). In contrast, the Explicitness Hierarchy presents a good format for diachronic change. Both for paradigms without a dual (see section 6.4) and for paradigms with a dual (see sections 9.3 and 9.4), this hierarchy is corroborated by the examples of cognate paradigms attested. In general, the diachronic dynamics of person paradigms are most clearly seen in the structure of the first person complex. The complete set of connections attested between the various types of marking for the first person complex is summarised in Figure 10.5. This map is a first outline of the similarity space of paradigmatic structure. The types that are directly connected by a line in the figure are strongly similar, as shown by the existence of cognate paradigms linking these types (see sections 6.5 and 9.6).



Figure 10.5: Connections attested between paradigmatic types by cognate paradigms

10.2 Prospects

10.2.1 Where to go from here

Even within the rather strong restrictions on the domain of the present investigation, the paradigmatic structure of person marking as attested in the world's languages has turned out to be highly variable. I believe to have presented a step in the direction of a better understanding of the marking of person in language. Yet, much work remains to be done. Most importantly, this investigation only addressed the PARADIGMATIC structure of person markers. The logical next step would now be to compare the SYNTAGMATIC function of these paradigms in a cross-linguistic investigation. The

paradigms that have been discussed in this study perform many different functions within the overall structure of a language. They mark, for example, the subject or the object of a predicate. Also, they can be more or less obligatorily required to form intelligible utterances in a language. These dimensions have not been separated in this investigation (cf de Groot & Limburg, 1986; Siewierska & Bakker, 1996). Finally, there might be large differences (or maybe rather interesting correspondences) between predicative uses of person markers (*I walk*) and other linguistic contexts, like isolated use (*it's me*) or pronominal possession (my book, cf Siewierska, 1998). All these questions can now be investigated with the present analysis of the paradigmatic structure as a possibly significant variable. The analysis of paradigmatic structure gives a proper basis for the investigation of the syntagmatic function.

Yet, also within the restrictions of the investigation into the paradigmatic structure of person marking, there are still many themes to be addressed. Two examples of the kinds of question that can be asked will be discussed in the next two sections. These sections address two observations that I have made during my investigations, but that could not be given a proper place within the main line of this work. The first observation concerns the place of the marking of gender oppositions within the structure of a paradigm of person. The second observations addresses a possible difference between prefixal and suffixal person marking. Both these themes deserve a proper and in-depth study.

10.2.2 Gender in person paradigms

As one of the main delimitations for the present study, I have discarded all paradigms that mark different forms for gender (see section 1.2.5). Still, I have had a more than incidental look at paradigms that mark gender, but this part of the investigation has not been sufficiently thorough to be reported on in the present study. Yet, there is one observation on the marking of gender in person paradigms that seems promising as a hypothesis for further research. Quickly formulated, gender marking is not found in paradigms that have an inclusive/exclusive opposition. However, gender marking has to be specifically defined for this hypothesis to be make sense.

When studying gender marking in the domain of person marking, two cases have to be distinguished. First, gender marking in the third person is a separate question altogether. Many semantic and formal classifications can be found here (Corbett, 1991). In contrast, natural gender (ie an opposition between male and female) is the only attested markedness strategy in the person markers that include at least speaker or addressee. Speaker and addressee are always human participants (or at least anthropomorphic non-humans). Apparently, the distinction between male and female is sufficiently dominant in the classification of human beings to appear time and again in the structure of human language. Gender marking in the reference to speaker or addressee can take many different forms. It is rather uncommon for reference to 'I', ie for reference to the speaker only. However, gender is rather often, cross-linguistically speaking, observed in the grammaticalisation of person markers like 'you' or 'we' (Plank & Schellinger, 1997). It is, for example, attested in Spanish, where two different forms for 'we' are found, *nosotros* for a group of males or mixed and *nosotras* for a group of females.

Now, the hypothesis can be precisely formulated. If there is gender marking in the reference to any person marker which includes reference to at least speaker or addressee, then such a paradigm will not also make a distinction between inclusive and exclusive ‘we’. In traditional terminology, if there is gender marking in the first or second person (singular or plural), there is no inclusive/exclusive opposition in the first person plural. From the data as reported on by Plank & Schellinger (1997), it can be inferred that the hypothesis fares rather well cross-linguistically. Counterexamples (ie paradigm that contain both an inclusive/exclusive opposition and a gender opposition in the first or second person) are almost unattested.⁶

If this hypothesis turns out to hold after further investigation, a possible explanation can be found in the concept of Pure Person. Paradigms that mark an opposition between inclusive and exclusive have restricted themselves to mark person, nothing else. Only if the marking of person is not completely explicit (ie there is no inclusive/exclusive opposition), is there a possibility for marking gender oppositions. One can look upon this as two different routes for the development of fine-grained referential categories for participant reference. On the one hand, a paradigm can specialise for the specific role the participants fulfil in the speech act. In this case, the paradigm marks ‘person’, and gender is not attested in the paradigm. Otherwise, a paradigm can specialise for the marking of intrinsic characteristics of the participant. In that case, the paradigm can mark gender. Yet, when gender is marked, the person-opposition ‘inclusive/exclusive’ is not attested.⁷

10.2.3 Asymmetry of affixation

Generalising over the world’s linguistic variation, there appears to be a strong preference for suffixation over prefixation. Grammaticalised elements predominantly are found after the root which they modify. For example, in a typological study, Bybee *et al.* (1990:4) found 426 grammatical prefixes (26%) against 1236 grammatical suffixes (74%). After a thorough investigation of the possible reasons for this predominance, the authors conclude that the preference for suffixation is due to the predominance of V-final languages

‘Our proposal for explaining the suffixing predominance, then is that grams at clause boundaries tend to affix at a very high rate, while the rate of affixation for clause-internal grams is determined by their meaning and relevance to the verb. The large number of suffixes in our sample, then, is due primarily to the fact that there are many more V-final language than V-initial, and the additional fact that the V-final languages are highly consistent in postposing verbal grams.’ (Bybee *et al.*, 1990:34).

⁶ The few cases that contradict the hypothesis are the Khoisan language !Xu, the Papuan languages Baniata and Vanimo and the Australian language Ndjébbana (Plank & Schellinger, 1997:74-77). Note that all these cases, except for !Xu, have a combination of gender and inclusive/exclusive marking in the dual forms. The occurrence of gender and inclusive/exclusive distinctions in the core of the person marking paradigms is an extremely exceptional structure for human language.

⁷ Probably, this gender-hypothesis can be extended to cover also honorific marking. Often it is even difficult to draw a clear dividing line between the marking of gender and honorifics. The hypothesis then is that paradigms that mark different honorific forms for speaker or addressee reference also do not mark an inclusive/exclusive opposition. As far as I know, the only exception to this hypothesis is the person marking of Acehnese (Durie, 1985:117). However, I have not dealt extensively with the marking of honorifics in the pronominal domain to be able to present a complete survey of the possible exceptions.

As for person markers, this predominance for suffixation is less clear. Reanalysing the data from Bybee *et al.* (1990: 9, 13, 15), there are 240 prefixed person markers (40%) against 354 suffixed person markers (60%) in their sample. There still is a preference for suffixation, although less strong than the overall preference. Moreover, Bybee *et al.* have counted each person marking morpheme individually. They do not correct for the fact that person markers in a language are found together in a paradigm. The person markers in a paradigm are normally either all prefixes or all suffixes. So it might be interesting to count paradigms instead of individual person markers.

In the data as presented in chapters 3 and 4 of the present work, there are 148 inflectional person marking paradigms (see section 5.8). Of these 148 inflectional paradigms, 72 cases are prefixal (49%) against 76 cases which are suffixal (51%). The difference between suffixation and prefixation has disappeared. However, the results from Bybee *et al.* are still correct. It turns out that the suffixal paradigms are generally larger; they contain more person markers. More specifically, there appears to be a strong correlation between the prefixal/suffixal status and the Horizontal Homophony Hierarchy. Paradigms without horizontal homophony are predominantly suffixing, while paradigms with horizontal homophony are predominantly prefixing (see Table 10.1). Why this should be the case remains a puzzle, the more so since the prefixal/suffixal difference does not seem to correlate with the Explicitness Hierarchy. There seems to be something specifically relevant in the phenomenon of horizontal homophony, supporting the occurrence of prefixation. Future research will have to investigate this correlation with more care and consideration than has been possible so far.

<i>1</i> → <i>Inclusive</i>	–	–	–	–	+		
<i>1</i> → <i>Exclusive</i>	–	–	–	+	+	others	
<i>2</i> → <i>2+3</i>	–	–	+	+	+		
<i>3</i> → <i>3+3</i>	–	+	+	+	+		Total
Number of inflectional paradigms	63	22	15	13	13	22	148
<i>Prefixal</i>	10	13	14	12	11	12	72
<i>Suffixal</i>	53	9	1	1	2	10	76
% Prefixal	16%	59%	93%	92%	85%	55%	49%

Table 10.1: Correlation between kind of affixation and horizontal homophony

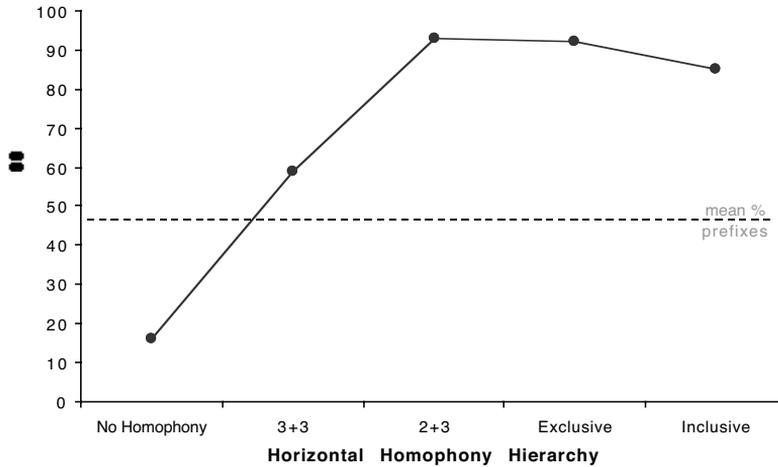


Figure 10.6: Percentage of prefixal paradigms on the Horizontal Homophony Hierarchy

10.3 Wider application of results

Cross-linguistic investigation and linguistic typology, the kind of research strategies that have been used in the presented investigation, do not stand in isolation from other fields of linguistic investigation. As has been argued throughout this work, the investigations build on a long linguistic tradition of analysis, description and classification. At the end now of my inquiries, the results of this study will hopefully be related to these domains of investigation and inspire future research.

One of the main insights of this study is that the ‘6-way’ paradigm –the kind of paradigm with three persons in the singular and in the plural– is in no way to be seen as the basic paradigmatic structure of person marking. It is indeed a rather commonly attested structure, but it is not the only common one. There are many other common paradigmatic structures. This insight can enrich both descriptive practice and historical reconstruction. The wide variation of paradigmatic structures that has been laid out in the preceding pages can guide the analysis of yet undescribed or only poorly described languages. Also, the typology of paradigmatic structure can be used in the context of a historical investigation to test whether a proposed reconstruction for a paradigm is in line with the attested distribution of diversity among the world’s languages. Of course, it is not necessary for a reconstructed paradigm to belong to one of the commonly found paradigmatic structures. However, the argumentation for the reconstruction of a typologically rare paradigmatic structure has to be much more stringent than the argumentation for a reconstruction of a paradigmatic structure that is commonly attested among the world’s languages. In this sense, the synchronic typology as presented in the present work can be used to evaluate historical reconstruction.

‘It has been proposed, as a basic contribution of synchronic typology to historical linguistics, that synchronic universals serve as touchstone for the validity or reconstructed systems.’ (Greenberg, 1969:174)

Finally, some insights have been presented that can inspire the development of the theory of linguistic structure. The regularity of the internal structure of paradigms of person, and especially the central place of the inclusive/exclusive opposition within this structure, are promising theoretical results. A more widely ranging prospect can be found in the notions variability and specialisation, as summarised in the concept of Pure Person. From the analyses presented of the paradigmatic structure of person marking, there appear to be clines between more and less explicit specialisation of a sub-system within the overall structure of a language. The more specialised a sub-system, the less variability is possible. This process can be interpreted as a kind of ‘meta’-grammaticalisation. Not only individual elements within a language can grammaticalise, but also sets of interconnected elements as a whole can grammaticalise. In such a process of paradigmatic grammaticalisation, changes in form and in function develop in an iconic relationship. Functionally, sub-systems of a language grammaticalise by streamlining the function of the morphemes included within the confines of a specific overall function and, formally, regularise the paradigmatic structure. However, at least in the case of the paradigmatic structure of person marking, it is doubtful whether this ‘meta’-grammaticalisation process is unidirectional. Developments in both directions seem possible.

Appendix A

Compound pronouns

Other person categories disqualified

A.1 Introduction

The category of person has been extended in the present work to include ‘general’ plural reference. It has been argued that the categories that are traditionally called ‘plural’ personal pronouns are better seen as a special form of ‘person’ marking, viz groups of persons. There is a claim in the literature that even more person categories are attested in the form of so-called ‘compound pronouns’ (Forchheimer, 1953: 132-135; Hagège, 1982: 112; Wiesemann, 1986a: viii; Noyer, 1992: 182-187). Compound pronouns are a special kind of compounds made up from two personal pronominal elements. In some languages such pronominal compounds are (semi) grammaticalised, which leads to rather extensive pronominal systems. Some authors propose to interpret these compound pronouns on a par with the singular and the plural pronouns, such as, for example, Voorhoeve:

‘It seems as if there are three contrastive levels in the pronominal system: singular simple pronouns contrasting to plural simple pronouns; and plural simple pronouns contrasting to [compound] pronouns.’ (Voorhoeve, 1967: 428)

When these compound pronouns are included as part of PERSON-domain, this would lead to large paradigms of person. For example, when simple and compound pronouns are taken together, then there are 18 different pronominal elements in Ghomala’, a Grassfields language in Cameroon. This leads Wiesemann to claim that Ghomala’ has ‘the most complex pronoun system as far as person categories are concerned’ (Wiesemann, 1986a: viii). However, it can be doubted whether these compound pronouns are really special PERSON markers. In this chapter, I will review the available descriptions of such compound pronouns, only to conclude that they are indeed a special kind of pronouns (and not normal compounds), but that the claim for special person categories is rather weak. Until more specific descriptions of these compound pronouns become available, it seems best not to include them in the person-domain.

These pronominal compounds are constructed in a very special way. The first part of the compound is a plural pronoun that incorporates the reference of the second pronoun. It looks as if the second pronoun is rather redundant. The second part of the compound is a specification of the group referred to by the first pronoun. This special way to form pronominal compounds is explained in section A.2. It will be argued there that these compounds are (semi) grammaticalised, and can reasonably be seen as

pronouns, not as compounds. Due to their special characteristics, I will refer to them as ‘compound pronouns’.¹ However interesting these compound pronouns may be, they are found only in a very restricted region of the world: all examples that I know of come from central-western Cameroon. These examples are discussed in section A.3. Also in that section, I present some general characteristics of these compound pronouns distilled from the rather diverse descriptions in the various grammars. Although the languages are closely related, both areally and genetically, some interesting variation in the construction of the compound pronouns is attested. Some generalisations over the compound pronouns from these languages are presented in section A.4.² It remains a question whether these compound pronouns are a viable cross-linguistic phenomenon, as they are only found in such a restricted region. However, there is an indication that it could have been more widespread, given a different linguistic world. It turns out that the way these compound pronouns are constructed is based on a phenomenon which is called the ‘incorporative reading’ of the first part of the compound (Hyman, 1979:53). This ‘incorporative reading’ of pronouns is widely attested in the languages of the world. A few examples of such ‘incorporative reading’ of pronominal reference are discussed in section A.5. This ‘incorporative reading’ is a precondition for the formation of compound pronouns. As this precondition is found widely dispersed in the world’s languages, I conclude that the compound pronouns are a real possibility of natural language. It may just be an incidental fact that they are only found in a very restricted region.³

¹ The oldest reference to these compounds as special pronouns is by Meinhof (1906:54), but he does not use a special name for them. Forchheimer (1953:132-135) follows up on the examples given by Meinhof. Forchheimer calls the elements ‘compound’ pronouns. They are also known as ‘complex’ pronouns. The first to use the term ‘complex’ was Voorhoeve (1967:427). In the recent descriptive literature, these pronouns are alternatively called ‘compound’ pronouns (eg Hyman, 1981:17; Fransen, 1995:183) or ‘complex’ pronouns (eg Hyman, 1979:53; Anderson, 1985:63; Parker, 1986:134). I will use the term ‘compound’ pronoun, as this name highlights the dual nature of these elements: partly specialised pronoun, partly compound of two pronouns.

² There are many more pronouns in the languages of the world that may be historically related to a compound of two pronouns. An obvious case is the inclusive, as, for example, *yumi* from Tok Pisin, which is composed of the second singular *yu* and the first singular *mi* (Foley, 1986:67). I will not discuss this kind of morphological compounds. There are some more examples of this sort discussed by Forchheimer (1953:98-12) and McGregor (1989:445). Forchheimer discusses both compounds that exist of a pronoun-pronoun combination, and compounds of different origins. If only those cases are considered where a pronoun originates from the combination of two pronouns, it turns out that only the inclusive ‘we’ is found as composed of two pronouns, ‘I + you’. Such compounds are interesting for a diachronic account of pronominal systems. However, for the present (synchronically oriented) work they are not of central importance. The main objective of this work is to develop an account of the different possible categories that can be found in pronominal paradigms. A compound like *yumi* from Tok Pisin does not add any new categories. The opposite is the case with the compound pronouns that are discussed in this chapter. They form a new kind pronouns, different from the possibilities that were discussed in the previous chapters.

³ The restricted distribution of compound pronouns can be compared with the situation of click-sounds, which are also only found in a few languages in a restricted region in the southern part of Africa. Clicks are nevertheless an option for human language, although a highly marked one.

A.2 From compound to pronoun

The idea of compound pronouns is exemplified here by the pronouns from Aghem, a Grassfields language from Cameroon. Aghem has seven monosyllabic simplex pronouns, a paradigm of the ‘Mandara’-type. The forms that are presented in (A.1) are the object forms, which are only slightly different from the subject forms (Hyman, 1979:47,49). The object forms are shown because there are used as a basis for the formation of the compound pronouns, to be discussed shortly.

(A.1) AGHEM

1	mùó	`sè´	1+2
			1+2+3
2	wò`	ghàʔ	1+3
3	`wín	ghě	2+3
		`ghé	3+3

Two such simple pronouns can be combined into a complex noun phrases, like *them and me*, shown in (A.2a). In English, a noun phrase like *them and me*, although syntactically possible, sounds pragmatically rather awkward. A sentence like *the photo shows them and me* can only be interpreted with a clear distance between the group of others, ‘them’, and the speaker, ‘me’. Pragmatically unmarked, one would rather use an exclusive pronoun, as in *the photo shows us*. Another possibility in English, with a slightly different meaning, is the sentence *the photo shows them with me*. However, in many languages, including Aghem, there is no special ‘and’ conjunction. These languages use the comitative preposition ‘with’ as a conjunctive between all noun phrases (Stassen, to appear). Consequently, there is only the possibility as shown in (A.2a), meaning something in between ‘them and me’ and ‘them with me’.

(A.2) AGHEM

a. *ghé à mùó*
 3+3 COM 1
 ‘them and me’ (Hyman, 1979:53)

b. *ghàʔ-à ghé*
 1+3-COM 3+3
 ‘us and them’, ‘me and them’, ‘us and him’ (Hyman, 1979:53)

At first sight, the phrase shown in (A.2b) is not much different: it shows a noun phrase meaning something like ‘us and them’. However, this phrase can also mean ‘me and them’ or ‘us and him’. Hyman explains about (A.2b):

‘all this [phrase] says is that there are at least three people and at least one of them has to be first person and another one has to be third person. It says nothing about the internal composition of the group. ... The two parts are welded together in a *incorporative* bond.’ (Hyman, 1979:53)

The term ‘incorporative’ refers to the fact that the reference of the second part of the compound is already incorporated in the first part. In (A.2b), the group 3+3, ‘them’, is a fraction of the group 1+3, ‘us’. This is different from both the English *us and them*, (where the groups are separate) and the English *us with them* (where the groups are joined into one group, but there is still no overlap between the referents of ‘us’ and

‘them’). These compounds seem to be most naturally translated into English by the first part only. Example (A.2b) means ‘us’, although the reference who are these persons is some more specified in Ghomala’. Compounds such as (A.2b), with an incorporative reading of the first pronoun, are considered to be a special structure, called COMPOUND PRONOUNS.⁴

There are nine such compound pronouns in Aghem. They are presented in (A.3). All compound pronouns consist of two simple (object) pronouns, with the linker ‘à’ incorporated morpho-phonetically into the first part. The compound pronouns are ordered in the table according to the FORM of the constituting parts: the forms with identical first parts are put in the same rows, the forms with identical second parts are put in the same columns. The constituent parts are identified by the numbers on the borders of the table. It can be seen that only a selected set of pronoun combinations are used as compound pronouns. It turns out that the first element of a compound pronoun is always plural. This is easily explained, as only plural pronouns can have an incorporative reading.

(A.3) AGHEM

	2	2+3	3	3+3
1+3	ghàʔà-wò	ghàʔà-ghè	ghàʔà-wìn°	ghàʔà-ghé
1+2				sàà-ghé
2+3			ghàà-wìn°	ghàà-ghé
3+3			ghèè-wìn°	ghèè-ghé

Combinations of pronouns in Aghem are either a ‘normal’ compound, as in example (A.2a), or a compound pronoun. Each combination of pronouns only occurs in one of both guises. The ‘normal’ compounds of two pronouns are called ‘cumulative’ by Hyman (1979:52). They are shown in Figure A.1. In this table, the dark-grey areas are those combinations that yield reciprocal or reflexive meanings – if they are interpretable at all. Reflexives and reciprocals are not included in this study, and consequently, these compounds are disregarded here.⁵ The light-grey area shows the combinations that form compound pronouns. Only in these combinations the incorporative reading of the first pronoun takes place. There is no overlap between the ‘incorporative’ compound pronouns and the ‘cumulative’ combinations of pronouns. There remain a few combinations that are not used. The meaning of these are taken care of by the reference of compound pronouns. It remains unclear, though, why only these are not attested as either a (cumulative) compound noun phrase or a (incorporative) compound pronoun.

⁴ The term ‘incorporative’ was coined by Hyman (1979:53). The opposite, as in example (A.2a) is called ‘cumulative’. Schaub (1985:198) uses the terms ‘selective’ and ‘additive’, respectively. I will follow Hyman’s terminology.

⁵ For Aghem there are indeed no compound pronouns for these combinations (Hyman, 1979:54), but this is not necessarily so. For another Grassfields language from Cameroon, Bamileke, reciprocal compound pronouns are described by Voorhoeve (1967:427).

	1	1+3	1+2	2	2+3	3	3+3
1				mùò à wò	mùò à ghě	mùò à wìn°	mùò à ghé
2	wò à mùò	wò à ghà?				wò à wìn°	wò à ghé
3	wìn à mùò	wìn à ghà?	wìn à sè	wìn à wò	wìn à ghě	–	
1+3							
1+2						–	
2+3	ghè à mùò					incorporative compound pronouns	
3+3	ghé à mùò	–	ghé à wò	–			

Figure A.1: Cumulative compound pronouns from Aghem

Also, there remain some questions about the meaning of the compound pronouns. This is partly because different descriptions describe the meaning of these pronouns rather differently, or hardly indicate the meaning at all. Some aspects are to be found in all descriptions of compound pronouns; I will start at this point of consensus. Basically, there seem to be two cases that should be distinguished: whether the second part of the compound pronouns is singular or plural. If the second part is a singular, then the compound pronoun has dual meaning. The pronoun *ghà?à-wò*, literally ‘we-exclusive with you-singular’, means ‘we including you’, or better: ‘I and you’. If the second part is plural, then either of the constituting persons can have plural meaning, or both, as exemplified above in (A.2b). So far so good. It begins to be unclear when the relation is examined between the ‘simplex’ pronouns and the compound pronouns. The difference is not one of reference. For example, the reference of the inclusive pronoun *sè* is roughly identical to the compound pronoun *ghà?à-wò*, and for that matter, also to the cumulative *wòàmùò*. They all refer to a group that consists of speaker and addressee. I suspect that there is a difference in the grouping of the participants between the various possibilities. Probably, a compound pronoun is used when there are two different groups set up in the discourse, that are now acting together for a short moment. The two groups are not put together into one new group, as they will remain separate later on. Parker (1986:136-137) hints at such an analysis. However, none of the descriptions gives a clear analysis of the use of the compound pronouns. The most direct analysis is found in a description by Voorhoeve (1967) on the compound pronouns (he calls them ‘complex pronouns’) of Bamileke, another Grassfields language from Cameroon. Voorhoeve states that the compound pronouns relate to the plural pronouns just as the plural pronouns relate to the singular ones.

‘The meaning of complex pronouns cannot be clearly distinguished from that of simple plural ones. ... However, the composition of the group of participants is much clearer defined, if complex pronouns are used. It seems as if there are three contrastive levels in the pronominal system: singular simple pronouns contrasting to plural simple pronouns; and plural simple pronouns contrasting to ... complex pronouns.’ (Voorhoeve, 1967:428)

In a sense, this makes a nice analysis: different singular participants can be combined into a group to form ‘plural’ pronouns. Then, different groups (whether singular or plural!) can be combined into a compound pronoun to form yet another kind of referential morpheme. Still, the fact remains that the reference of plural pronouns and

compound pronouns overlap. A detailed analysis is needed to decide whether compound pronouns are indeed a further level of group-formation, or whether they are just a pragmatical side-effect of compounding pronouns. For now, their status as ‘pronouns’ remains doubtful, although promising.

To summarise, there are four characteristics of compound pronouns. The first characteristic is closest to a definitional property: the reference of this first plural pronoun INCORPORATES the reference of the second part. The reference of the individual parts of the compound pronoun are not to be added together; the second part is only a clarification of the first part. Consequently, there are often different referential possibilities. A compound pronoun can, for example, refer to ‘we and he’, ‘we and they’ or ‘I and they’. Second, the linker IS FUSED morphophonetically into the first part, or both parts of the compound are fused. This indicates grammaticalisation of the compound. Third, the first part of the compound is always a PLURAL pronoun – this characteristic in fact follows from the first. For an incorporative reading, a non-singular pronoun is necessary. And finally, the compound pronouns do not mark anything that is not possible to say with the simple pronouns. The compound pronouns are probably PRAGMATICALLY MARKED elements, but they do not seem to have a special referential value. The first two characteristics – incorporative meaning and fusion – are the strongest indications that there is indeed a special phenomenon here. It is noteworthy, though, that it is not found more widespread among the languages of the world. The only descriptions of compound pronouns come from languages that are spoken in a very restricted part of Cameroon, to be discussed in the next section. However, the incorporative reading of plural pronouns occurs more often, also when there are no grammaticalised compound pronouns in the language. A few examples of the incorporative reading outside Cameroon will be discussed in section A.5.

A.3 The Bantoid compound pronouns

There are numerous descriptions of compound pronouns found in grammars of Bantoid languages from Cameroon. All these descriptions take a slightly different view on the phenomenon, which makes it interesting to combine the insights of the different grammarians. I will discuss the structure of the compound pronouns from the languages mentioned in (A.4), all member of the Bantoid subgroup of the Niger-Congo family. Most examples of compound pronouns are found among the Grassfields subgroup of the Bantoid languages. All languages in (A.4) are spoken in Western Cameroon, on the border with Nigeria (see the map in Grimes, 1996: 188). It is a clearly areal phenomenon.

- (A.4) Bantoid: – Beboïd: **Noni**
 – Narrow Bantu: **Nkosi**
 – Grassfields: – Mbam-Nkan: **Bamileke, Ghomala’,
 Ngiemboon, Limbum**
 – Momo: **Mundani**
 – Ring: **Babungo, Aghem**

The pronouns and compound pronouns in these languages will be discussed in the following subsections. These subsections are placed in ascending order according to the number of compound pronouns found in the language, starting from 4, with a maximum of 10. In each subsection I will rephrase the description of the compound pronouns as it is found in the grammars. To ensure comparability, I will analyse all pronominal systems according to a few principles. First, I will split the systems in monosyllabic forms and polysyllabic forms. The monosyllabic forms will be the simplex pronouns, and the polysyllabic forms will be the compounds.⁶ The polysyllabic compound pronouns will be analysed according to the parts they are made of. In almost all cases, the compound pronouns are transparent combinations of simplex pronouns. When the compound pronouns are ordered according to the constituting parts, then it appears that the compound systems are structured highly alike, something which is not directly obvious from the individual descriptions.⁷

Babungo

The simplex – non-compound – pronouns from Babungo are shown in (A.5). Babungo has a set of 8 monosyllabic argument pronouns (Schaub, 1985: 193-194).

(A.5) BABUNGO

		nsôo	1+2
		nsíŋ	1+2+3
1	mə̀	yia	1+3
2	ghɔ̀	vɪŋ	2+3
3	ŋwé	vǎŋ	3+3

Babungo has the simplest set of compound pronouns of all languages discussed here. There are four compound pronouns that are made by compounding the subject pronouns.⁸ They are shown in (A.6). The first part is a plural first or second person, that has an incorporative reading (called ‘selective’ in the grammar). The second part is either a singular or a plural third person (Schaub, 1985: 198).⁹

(A.6) BABUNGO

		3	3+3
1+3	yía-ŋwé	yía-vǎŋ	
2+3	víŋ-ŋwé	víŋ-vǎŋ	

⁶ I will deviate from this principle slightly in the cases of Limbum and of Noni.

⁷ I will disregard all forms that have logophoric reference (Hagège, 1982: 105). Special logophoric third person forms are found regularly in this part of Africa, and in some languages they also take part in the formation of compound pronouns, eg Noni (Hyman, 1981: 15-18) and Babungo (Schaub, 1985: 198).

⁸ The compound pronoun *víŋ-yi* with a logophoric third person as second part is disregarded here.

⁹ There are a few more compound pronouns mentioned in the grammar. These, however, all have obligatory plural reference of the first part. This is comparable to the cumulative compound pronouns as described in section 2 (they are called ‘additive’ in the grammar). The cumulative compound pronouns are made with a linker *vi*, an indefinite pronoun (Schaub, 1985: 196, 198). The form *yíavìvǎŋ* means either ‘We and he’ or ‘we and they’. There are only three different cumulative compounds: *yíavìvǎŋ*, *nsíŋvìvǎŋ* and *víŋvìvǎŋ*.

Mundani

Relative to Babungo, the number of compound pronouns is doubled in Mundani. However, there are fewer non-compound subject pronouns, as shown in (A.7). Mundani has only a 6-way monosyllabic pronoun system (Parker, 1986: 132).¹⁰

(A.7) MUNDANI

1	má	pá	1+2
		b̄ĩ	1+2+3
2	à	b̄ĩ	1+3
3	tà	b̄ò	2+3
		b̄ò	3+3

The first part of the compound pronouns, as shown in (A.8), are the plural pronouns; the second part is more complicated. The forms *nè* and *n̄ĩ*, in the first two columns, are second person forms that are normally found only after the preposition *ne*, ‘with’. In the second and fourth column (the second part of the compound is plural here) a pluraliser *bá* is added (Parker, 1986: 135).

(A.8) MUNDANI

	2	2+3+ <i>Plur</i>	3	3+ <i>Plur</i>
1+3	bá nè	bá n̄ĩ bá	bá t̄ò	bá tò bá
2+3			bí t̄ò	bí tò bá
3+3			bó t̄ò	bó tò bá

Ngiemboon

The situation in Ngiemboon shows a possible transition between compound and simplex pronouns. It looks as if there is an 8-way pronoun paradigm, shown in (A.9), although the inclusive forms 1+2 and 1+2+3 are not monosyllabic (Anderson, 1985: 63).¹¹

(A.9) NGIEMBOON

		pógò	1+2
		pégè	1+2+3
1	mèŋ	pég	1+3
2	gù/ð	pĩ	2+3
3	yé/à	pó(b)	3+3

The two polysyllabic pronouns *pógò* and *pégè* are probably fused compound pronouns (Anderson, 1985: 63). When they are still taken as compound pronouns in origin, the Ngiemboon system of compound pronouns, shown in (A.10), is identical to the Mundani system of compound pronouns. A compound meaning ‘we and you’ can obviously grammaticalise into an inclusive pronoun.

¹⁰ I was unable to reproduce the exact tonal orthography of the original source with the fonts that I have a my disposal. The differences are kept as minimal as possible.

¹¹ The paradigm presented in (A.9) consists of the morphemes that are used as subject. There are slightly different morphemes used for the object. The object form of the compound pronouns is made from the simplex object pronouns (Anderson, 1985: 68).

(A.10) NGIEMBOON

	2	2+3	3	3+3
1+3	pêg+ɔ > pógò	pêg+a+pi > pégè	pèg yè	pég-à pò
2+3			pi yè	pí-a pò
3+3			pò yè	pób-à pǒ

Nkosi

Almost the same situation as in Ngiemboon is also found in Nkosi (Dorsch, 1911:249-250). There are 8 monosyllabic pronouns, shown in (A.11). The two inclusive pronouns are probably of compound origin. Dorsch (1911:250) comments that the minimal inclusive (1+2) *sū* comes from the compound *se* + *ǒé*. The origin of the augmented inclusive (1+2+3) *séa* is not explained. It is possibly a contraction of the still existing compound pronoun *seanyi*.

(A.11) NKOSI

		sū	1+2
		séa	1+2+3
1	me	se	1+3
2	ǒé	nyí	2+3
3	mò	bò	3+3

Besides the fused compound *sū*, there are 7 clear compound pronouns, shown in (A.12). If the pronoun *sū* is counted as a compound pronoun, then this paradigm has the same structure as the Ngiemboon system. Whether there is a difference in meaning between the simplex pronoun *séa* and the compound pronoun *seanyi* is unclear. Dorsch (1911:250) translates *séa* as ‘we and you-plural’ and *seanyi* as ‘I and you-plural’. This seemingly clear difference is not corroborated by the analyses of the same form in other languages. In other languages, the corresponding form to *seanyi* is normally translated as ‘we and you-plural’ (eg Hyman, 1979:53).

(A.12) NKOSI

	2	2+3	3	3+3
1+3	se+oe > sū	se-a-nyi	su-mò	se-a-bò
2+3			nyi-mò	nyi-a-bò
3+3			bu-mò	bo-bò

Bamileke

A slight variation on the last examples can be found in Bamileke. This time, there are seven monosyllabic pronouns, as shown in (A.13). The inclusive *bàn* belongs in this paradigm, although it shows slightly different tone-behaviour (Voorhoeve, 1967:422).¹²

¹² Shown here are the simplex object pronouns. The simplex subject pronouns differ slightly from these. The simplex object pronouns are shown as they are used to make the compound pronouns, also when the compound is a subject.

(A.13) BAMILEKE

		b̀̀n	1+2
		°bag'	1+2+3
1	mə	°bin'	1+3
2	ui	bó	2+3
3	jé		3+3

The system of compound pronouns, shown in (A.14), is structurally like the Ngiemboon system. Interestingly, there is already an inclusive form in the simplex system, which seems to make the 'we and you' compound pronouns superfluous (cf. the occurrence of *séa* and *seanyi* in Nkosi). Still, both forms exist. Probably *bâg-ù̀̀* and *bâg-à-bin`* are used with a different pragmatical value as the inclusive pronoun *b̀̀n* (Voorhoeve, 1967:427).

(A.14) BAMILEKE

	2	2+3	3	3+3
1+3	bâg-ù̀̀	bâg-à-bin`	bâg-jé	băg-à-bo
2+3			bîn-jé	bín-à-bo
3+3			bô-jé	bó-à-bo

Limbum

Limbum has an 8-way simplex pronoun system (Fransen, 1995:179-180). The two inclusive forms, as shown in (A.15), do not show a strong indication that they are composed of two parts. Only the fact that *sì̀̀* and *wì̀̀* end identically in *ì̀̀* could indicate a historical relation between those forms.

(A.15) LIMBUM

		sò	1+2
		sì̀̀	1+2+3
1	mè	wìr	1+3
2	wè	wìi	2+3
3	(y)í	wōwì	3+3

The compound pronouns, as shown in (A.16), are rather unusual (Fransen, 1995: 183-185). The most interesting feature are the first two columns (A.16). The two compound pronouns in these columns are composed of a second person plural with the two simplex inclusions. This kind of compound is not found in any other language with compound pronouns. These compounds mean the same as the combination first person plural/second person, like in Ngiemboon and in most languages discussed here. The meaning of the third column of the complex pronouns (with *-yì* as second part) is slightly different in Limbum, when compared to the other languages cited here. In all other languages these forms are strictly dual. The plural first part of the compound can only have the incorporative reading. In Limbum, however, the compounds shown here in the third column can also be used with reference to more than two persons, meaning respectively 'we and he/she', 'you (plural) and he/she' and

‘they and he/she’ (Fransen, 1995:184). It is unclear why Limbum deviates from the other languages in this aspect.¹³

(A.16) LIMBUM

	1+2	1+2+3	3	3+3
1+3			wìr-yī	wìr-wōwī
2+3	wìl-sò	wìl-siì	wìl-yī	wìl-wōwī
<i>concord</i>			ó-yī	ó-wōwī

The use of the simplex inclusive pronouns as part of the compound pronouns in Limbum indicates that these inclusive pronouns are strongly grammaticalised. They are real parts of the simplex pronominal system. This is quite different from the situation in Mundani, Ngiemboon and Nkosi, where the inclusives are still overtly related to the compound pronouns.

Noni

The simplex pronouns from Noni are shown in (A.17). It is a 7-way system with a specialised inclusive (Hyman, 1981: 15).

(A.17) NONI

		beènè	1+2
		bèsèn	1+2+3
1	me	bèn	1+3
2	wò	bò	2+3
3	wvù	bó	3+3

The compound pronouns are shown in (A.18). Note that the 1+3 part of the compound pronouns, *beè*, is different from the simplex 1+3, *bèsèn* (Hyman, 1981:17). Historically, they are probably based on the same element, maybe a compound. Hyman analyses the simplex inclusive pronoun *beènè* as a compound pronoun, historically derived from a combination of the two pronouns *beè*+*bèn* (Hyman, 1981: 15, 17). The resulting pronoun *beènè*, however, is grammaticalised as a simplex form. It is used in turn as the first part of the compound *bènè-bò*.

(A.18) NONI

	2	2+3	3	3+3
1+3	beè-wò	beè+bèn > beènè	beè-wvù	beè-bò
1+2				beènè-bò
2+3				bènè-wvù
3+3				bò-bò

Aghem

For comparison with the other languages, I repeat the data from Aghem here from section A.2. There is a 7-way simplex pronoun system, just as in Noni (Hyman, 1979:49).¹⁴

¹³ Another special aspect of the Limbum compound pronouns is the use of the concord-marker *ó-* instead of the third person plural (Fransen, 1995:193).

(A.19) AGHEM

		\`sè´	1+2
			1+2+3
1	mù´	ghàʔ´	1+3
2	wò´	ghè´	2+3
3	\`wìn	\`ghé	3+3

The compound pronouns resemble the structure from Noni strongly. The only difference with Noni is that the compound 1+3 with 2+3 is not related to the simplex inclusive '\`sè´'. Aghem thus has a true 9-way compound pronoun system.

(A.20) AGHEM

	2	2+3	3	3+3
1+3	ghàʔà-wò	ghàʔà-ghè	ghàʔà-wìn°	ghàʔà-ghé
1+2				sàà-ghé
2+3			ghàà-wìn°	ghàà-ghé
3+3			ghèè-wìn°	ghèè-ghé

Ghomala´

Finally, there is Ghomala´, which has the largest inventory of compound pronouns that I know of. This extensive set of pronouns tempted (Wiesemann, 1986a:viii) to claim that Ghomala´ has the world's most extensive pronouns system. Unfortunately, the short note by Wiesemann is the only information available on this language. I will analyse these data within the framework that has been developed in this chapter. Still, there will be numerous questions that remain to be answered on the composition of the pronouns if and when more information becomes available.

The monosyllabic pronouns from the short note by Wiesemann (1986a:viii, citing Fossouo) form an 8-way person system. The three forms for 'we' – *pu*, *pə* and *pyə* – show a strong phonological similarity. This makes it probable that both inclusive forms have a compound origin, like in Ngiemboon. However, the augmented inclusive (1+2+3) *pə* is a constituent part of the compound pronouns, like the inclusive in Noni. This shows that the augmented inclusive pronoun is now completely grammaticalised as a simplex pronoun.

(A.21) GHOMALA´

		pu	1+2
		pə	1+2+3
1	gó	pyə	1+3
2	o	po	2+3
3	e	wap	3+3

[continued from previous page]

¹⁴ Shown here are the simplex object pronouns. The reason is identical to the one given for Bamileke (see footnote 12): the object pronouns are used for the construction of the compound pronouns.

The 10 compound pronouns exhibit the same basic combinations that were found in all the other languages: a plural first part with a third person second part. The third plural morpheme *-pu-* is not found among the monosyllabic forms in the table above, but it is clear from the meaning that it is a third person plural. The etymological origin of the element is unknown to me. The homophony with the monosyllabic 1+2 *pu* is probably a coincidence. Note that there is no *-a-* linker when the second part is a singular third person (first column). This linker is only found when there is a plural third person as second part (second column). The last two columns of (A.22) remain problematic. It is unclear what the forms mean precisely, nor is explained what the origin of the second part is. The use of the *-a-* linker in all these forms indicates a plural second part.

(A.22) GHOMALA'

	3	3+3(?)	(?)	(?)
1+3	pyə-é	pyə-a-pu	py-a-yú	py-a-wu
2+3	po-é	po-a-pu	po-a-yú	
3+3(?)	pu-é			
1+2+3		pə-a-pu	pə-a-yú	

The number of compound pronouns in Ghomala' is indeed unrivalled by the other systems that have been discussed in this section. Whether this makes it really the world's most complex pronoun system, as Wiesemann claims, depends on the choice to consider compound pronouns as pronouns or not.

A.4 Generalisations

After the individual languages have been discussed, a few generalisation on the structure and meaning of compound pronouns can be made. First I will make some generalisations over the form of the compound pronouns, followed by some observations on their meaning or referential value.

First, the STRUCTURE of the compound pronoun systems is very similar. However, although there is a strong similarity between the various paradigmatic structures, there are also strong indications that compound pronouns are an individually acquired phenomenon, and not a genetic feature. In all languages, the compound pronouns are constructed out of the simplex pronouns of the particular language. Although the simplex pronouns might be historically related between the languages, the compound pronouns are clearly made up from the 'own' simplex pronouns of a language. The compound pronouns themselves are not directly related to the compound pronouns of other languages. There are some differences between the languages as to which kind of compound pronouns are used, and as to how many different compounds are distinguished. However, there are a few combinations that are found in all languages. The kind of compound construction that is commonly found in compound pronouns is shown in Table A.1. Although these combinations are found in all languages, still, these compound pronouns in each language are related to the simplex pronouns of that same language, not to the structurally equivalent compound pronouns of other languages. From these characteristics, it appears that compound pronouns are an ex-

ample of structural diffusion. The same principle is found in all languages, but the languages all use their own inventory to make the compound pronouns. The range of different compounds that are used, is to be decided by each language alone. The idea to make compound pronouns is probably borrowed (or developed parallel), not the compound pronouns themselves.

1+3 compounded with 3	1+3 compounded with 3+3
2+3 compounded with 3	2+3 compounded with 3+3

Table A.1: Compound pronouns that occur in all discussed languages

Another structural characteristic is the linker-element *-a-*. In many languages that were discussed, this linker is used between the parts of the compound. In central Africa, this *-a-* is a widespread morpheme to link linguistic elements. Interestingly in this context, the linker *-a-* is mainly found with compound pronouns that have a plural second part. The only exception to this generalisation is found in Aghem.

Second, the MEANING of the compound pronouns also turns out to be strongly similar. When the second part of the compound is originally a singular pronoun, then the compound has dual meaning. A compound of the form ‘we and he’ means ‘I and he’. The only exception to this generalisation is found in Limbum, where the compound with the form ‘we and he’ can both mean ‘I and he’ and ‘we and he’. When the second part is plural, then both parts of the compound can have a plural reference. A compound with the form ‘we and they’ can either mean ‘I and they’, ‘we and he’ or ‘we and they’.

Finally, there is a strong referential connection between compound pronouns meaning ‘I and you’ and inclusive pronouns (1+2, 1+2+3). In some of the languages, compound pronouns with the meaning ‘I and you’ are found to grammaticalise into simplex inclusive pronouns. Once they are completely grammaticalised, these simplex inclusive pronouns can again be used as building part for complex pronouns. Different stages in this grammaticalisation are documented by the structures found in Ngiemboon, Nkosi, Noni and Aghem (in that order).

A.5 The incorporative plural revisited

The central property of compound pronouns as found in the Bantoid languages, is the peculiar INCORPORATIVE reading of the first pronoun.¹⁵ The first part of a compound pronoun is taken as a reference to the whole, incorporating the reference of the second part. A phrase that literally says ‘we and you’ means rather something like ‘I and you’. Although compound pronouns are only found in a restricted area in Cameroon, the incorporative use of a plural pronoun can be found widespread in the languages of the world.

¹⁵ This is ‘peculiar’, of course, only from a European point of view, not from the point of view of those language that have it.

Some languages use a plural pronoun where one would expect a singular one in conjunctions. For example in Bari, a Nilotic language from Sudan, when pronouns are conjoined, the first pronoun always has to be plural, also when only a single participant is meant. This is exemplified in (A.23 a,b). However, there are no complex pronouns in Bari, at least, they have never been described as such.¹⁶

‘When the conjunctive preposition *ko* joins pronouns, the pronoun preceding it must always be *plural*, no matter whether it stands for a single noun. ... In this way there is no telling whether two or more persons are concerned.’ (Spagnolo, 1933:212-213)

(A.23) BARI

a. *yi ko do*
 1+3 with 2
 ‘I and you’ (lit. ‘we with you’) (Spagnolo, 1933:212)

b. *ta ko nan*
 2+3 with 1
 ‘You and I’ (lit. ‘you all with me’) (Spagnolo, 1933:212)

Other examples, outside Africa, can be found in Mundari, a Munda language from India, and Tagalog, an Austronesian language from the Philippines. Both are examples of a conjunction of a pronoun with a proper name, not of two pronouns. Still, the pronoun in both cases has the incorporative reading. The Mundari example in (A.24) can only mean ‘I and Paku’, because the pronoun is dual. The incorporative reading of this dual pronoun implies that the total group consists of two persons, one of which is Paku, and consequently the only other is the speaker. The Tagalog example in (A.25), using a plural pronoun, can mean both ‘them and Juan’ and ‘him and Juan’.

(A.24) MUNDARI

aling Paku-lo
 1+3DL Paku-COM
 ‘I and Paku’ (lit. ‘we with Paku’) (Hoffmann, 1903:24-25)

(A.25) TAGALOG

sila ni Juan
 3+3 CONJ Juan
 ‘them/him and Juan’ (Schachter & Otones, 1972:116)

A last example of an incorporative reading of a plural pronominal marking comes from West Greenlandic Inuktitut, an Eskimo-Aleut language from Greenland. Here it is the inflectional pronominal marking that gets an incorporative reading, as shown in (A.26). The literal translation may sound like normal English, but there is an interesting difference between the Inuktitut and the English meaning. In the Inuktitut sentence, there are only two persons who will leave, so the best translation is ‘You and I will leave’. The literal translation in English, ‘we will leave with you’, implies that

¹⁶ Meinhof (1906:54) mentions Duala, a Bantu language from Cameroon, as an example with the same pronoun use. He claims also that this is a widespread phenomenon among Bantu languages. It is difficult to judge whether compound pronouns really are only found in Cameroon, or whether this strong areal restriction is a result of descriptive practice. Maybe compound pronouns is a phenomenon only known to specialists in the languages in and around Cameroon, and stays unnoticed in the description of other African languages.

there will be more than two, as the ‘you’ is normally not interpreted as part of the ‘we’. The Inuktitut ‘we’ inflection has an incorporative reading.

(A.26) INUKTITUT

illil-lu aalla-ssa-aqut

2-COM leave-FUT-1+3

‘You and I will leave’ (lit. ‘we will leave with you’) (Fortescue, 1984:257)

These examples can only indicate that the incorporative use of the plural pronouns is probably widespread among the languages of the world. The examples in this section are from all over the world (Bari in Africa, Mundari in Asia, Tagalog in the Pacific and Inuktitut in America) and they are genetically diverse. Compound pronouns can best be seen as (semi)grammaticalised use of the incorporative use of plural pronouns. Probably, if compound pronouns would be a wider known phenomenon, they would be described for other languages as well.

A.6 Conclusion

Pronouns or compounds, that is the question. I have argued in this appendix that the compound pronouns of the Bantoid languages in Cameroon are to be included as a special kind of ‘pronouns’. The compound pronouns show a rather different structure from normal compounds: they are clearly grammaticalised in comparison to normal compounds, and the semantic interpretation is different from other compounds. This semantic interpretation is the most salient characteristic of these pronouns. The first part of the compound pronouns is to be interpreted ‘incorporative’: it includes the reference of the second part of the pronoun. This ‘incorporative’ reading of a pronoun is found more widespread among the world’s languages, indicating that the restricted occurrence of compound pronouns (it is only found in Cameroon) is probably only incidental.

A different question is whether these compound pronouns are new person categories, leading to the largest inventory of person categories among pronominal systems, as claimed by Wiesemann (1986a: viii). I doubt this. The referential value of these compound pronouns is often identical to the simplex group pronouns. However, more in-depth research is needed on the use of the compound pronouns as opposed to the group pronouns. If there are contexts which require a compound pronoun, and which do not allow for the simplex pronouns to be used, that would be an argument for a special referential status, although restricted contextually. From the available sources, it appears that the compound pronouns are freely interchangeable with simplex group pronouns, yielding a different pragmatical value, not a different referential value. Consequently, until further notice, the compound pronouns are not included as special person categories.

Paradigmatic variation with 3 forms for ‘we’

Paradigm	Examples	Page	To	S I	S W	S D	V H	I In	1 Ex	2	3
Dual-in-exclusive-only	Yagua, Ngankikurungkurr, Savosavo	251	3	-	+	+		+	-	-	-
Dual-in-inclusive-only	Middle Paman sub-family (Wik-Mungkan)	253	1	-	+	+		-	+	-	-
Inc/Exc-in-dual-only	Samo, Coos	253	2	-	+	+		-	-	-	-
Inc/Exc-in-plural-only	Kuku-Yalanji, Jiarong, Guhu-Samane, Tuaripi	254	4	-	-	+		-	-	-	-
Korafe-type	Korafe	256	1	-	-	-	2/3	-	-	-	-
Burarra-type (1)	Burarra	256	1	+	-	+		-	-	-	-
Total			12								

Paradigmatic variation with 4 forms for ‘we’

Paradigm	Examples	Page	To	S I	S W	S D	V H	I In	1 Ex	2	3
dual-inclusive/exclusive	Oceanic (Maori, Samoan), Pama- Nyungan (Dhalandji, Gumbaynggir), Tibeto-Burman (Bahing, Bunan), Chinook, Siuslaw, Iroquoian (Oneida, Cherokee), Mazehua, Kunama	257	12	-	+	+		-	-	-	-
Limbu-type	Tibeto-Burman (Limbu, Camling)	259	2	-	+	+		-	-	-	+
Dhuwal-type	Dhuwal, Sedang	259	2	-	+	+		-	-	+	-
Kilivila-type	Kilivila	260	1	-	+	+		-	-	+	+
Rapanui-type	Rapanui	260	1	-	+	+		-	-	+	+
Suana-type	Suana	261	1	-	+	-	2/3	-	-	-	-
Kwamera-type	Kwamera	261	1	-	+	-	In/3	-	+	+	-
partial-unit-augmented	Umpila, Alawa, Warrwa, Ngalakan, Maranungku, Jaminjung, Apinayé	262	7	+	+	+		-	-	-	-
Total			27								

Paradigmatic variation with 5 forms for ‘we’

Paradigm	Examples	Page	To	S I	S W	S D	V H	I In	1 Ex	2	3
unit-augmented	Gunwinguan (Rembarrnga, Ngandi), Djeebanna, Weri, Reefs, Kayapo	264	6	+	+	+		-	-	-	-
Burarra-type (2)	Burarra	266	1	+	+	-	Ex/2	-	-	-	-
Kunimaipa-type (3)	Kunimaipa	267	1	+	+	+		-	-	-	-
Total			8								

Paradigmatic variation with duals outside the first person complex

Paradigm	Examples	Page	To	S I	S W	S D	V H	I In	1 Ex	2	3
Dizi-type	Dizi	268	1	-	-	+		+	+	-	+
Achumawi-type	Achumawi	268	1	-	+	+		+	+	+	-
Tunica-type	Tunica, Slovene, Buin	268	3	-	-	+		+	+	+	-
Aleut-type (2)	Aleut	269	1	-	-	+		+	+	-	-
Total			6								

Paradigmatic variation with numbers higher than the dual

Paradigm	Examples	Page	To
Kiwai-type	Kiwai	234	1
Ambrym/Paamese-type	Oceanic (Ambrym, Paamese)	234	2
Sursurunga-type	Sursurunga	237	1
Yimas-type	Yimas	237	1
Biak-type	Biak	238	1
Total			6

References

- ABED, SHUKRI B. (1991). *Aristotelian Logic and the Arabic Language in Alfarabi*. Albany: State University of New York.
- ABONDOLO, DANIEL (1998a). “Khanti”. In: Daniel Abondolo (ed.) *The Uralic Languages*, pp. 358-386. London: Routledge.
- (ed.) (1998b). *The Uralic Languages*. London: Routledge.
- ADELAAR, WILLEM F.H. (1977). *Tarma Quechua: grammar, texts, dictionary*. Ph.D. dissertation, Universiteit van Amsterdam.
- (1993). “La Categoría del Plural Inclusivo en las Lenguas Americanas: Interpretación y Realidad”. In: Maxim P.A.M. Kerkhof, Hugo de Schepper & Otto Zwartjes (eds.) *España: ¿Ruptura 1492?*, pp. 207-216. (Diálogos Hispánicos, 11). Amsterdam: Rodopi.
- ADELAAR, WILLEM F.H. & LINCOLN SILVA LÓPEZ (1988). “Grammaticaal overzicht van het Guarani”. *Wampum* 7: 11-60.
- AIKHENVALD, ALEXANDRA A.Y. (1995). *Bare*. (Languages of the World/Materials, 100). München: Lincom.
- (1998). “Warekena”. In: Desmond C. Derbyshire & Geoffrey K. Pullum (eds.) *Handbook of Amazonian Languages*. Vol. 4, pp. 225-440. Berlin: Mouton de Gruyter.
- ALLEN, EDWARD J. (1976). “Dizi”. In: M. Lionel Bender (ed.) *The Non-Semitic Languages of Ethiopia*, pp. 377-392. East Lansing: African Studies Center.
- AMAYA, MARIA TRILLOS (1999). *Damana*. (Languages of the World/Materials, 207). München: Lincom.
- ANCEAUX, J.C. (1965). *The Nimboran Language: Phonology and Morphology*. (Verhandelingen van het Koninklijk Instituut voor Taal-, Land- en Volkenkunde, 44). Den Haag: Martinus Nijhoff.
- ANDERSEN, HENNING (1998). “Slavic”. In: Anna Giacalone Ramat & Paolo Ramat (eds.) *The Indo-European Languages*, pp. 415-453. London: Routledge.
- ANDERSEN, TORBEN (1988). “Ergativity in Pari, A Nilotic OVS Language”. *Lingua* 75: 289-324.
- ANDERSON, LAMBERT (1966). “The structure and distribution of Ticuna independent clauses”. *Linguistics* 20: 5-30.
- ANDERSON, LLOYD B. (1982). “The ‘Perfect’ as a Universal and as a Language-Particular Category”. In: Paul J. Hopper (ed.) *Tense-Aspect: Between Semantics & Pragmatics*, pp. 227-264. (Typological Studies in Language, 1). Amsterdam: John Benjamins.
- ANDERSON, STEPHEN C. (1985). “Animate and Inanimate Pronominal Systems in Ngyemboon-Bamileke”. *Journal of West African Languages* 15 (2): 61-74.
- DE ANGULO, JAIME & L.S. FREELAND (1931). “The Achumawi Language”. *International Journal of American Linguistics* 6: 77-120.
- ASHER, R.E. (1982). *Tamil*. (Lingua Descriptive Studies, 7). Amsterdam: North Holland Publishing Company.
- ASHER, R.E. & T.C. KUMARI (1997). *Malayalam*. (Routledge Descriptive Grammars). London: Routledge.
- AUSTERLITZ, ROBERT (1959). “Gilyak”. *Word* 15: 102-109.

- AUSTIN, PETER & JOAN BRESNAN (1996). "Non-configurationality in Australian Aboriginal Languages". *Natural Language and Linguistic Theory* 14: 215-268.
- AUSTING, JOHN & RANDOLPH UPIA (1975). "Highlights of Ömie Morphology". In: Tom E. Dutton (ed.) *Studies in Language of Central and South-East Papua*, pp. 513-598. (Pacific Linguistics, C 29). Canberra: Australian National University.
- BAKER, MARK (1990). "Pronominal Inflection and the Morphology-Syntax interface". *Chicago Linguistic Society* 26 (1): 25-48.
- (1991). "On Some Subject/Object Non-Asymmetries in Mohawk". *Natural Language and Linguistic Theory* 9: 537-576.
- BAUMAN, JAMES JOHN (1975). *Pronouns and Pronominal Morphology in Tibeto-Burman*. Ph.D. dissertation, University of California, Berkeley.
- BAVIN, EDITH L. (1981). "Lango: some Morphological Changes in Verb paradigm". In: Thilo C. Schadeberg & M. Lionel Bender (eds.) *Nilo-Saharan, proceedings of the First Nilo-Saharan Linguistic Colloquium*, pp. 89-100. Dordrecht: Foris.
- BEE, D. (1973). "Notes on Auyana Phonology and Morphology". In: Howard P. McKaughan (ed.) *The Languages of the Eastern Family of the East New Guinea Highland Stock*, pp. 186-189. Seattle: University of Washington Press.
- BENDER, LIONEL M. (1996). *Kunama*. (Languages of the World/Materials, 59). München: Lincom.
- BENDOR-SAMUEL, DAVID (1972). *Hierarchical Structures in Guajajara*. (Summer Institute of Linguistics Publications in Linguistics and Related Fields, 37). Oklahoma: University of Oklahoma.
- BENVENISTE, ÉMILE (1966). *Problèmes de linguistique générale*. (Bibliothèque des sciences humaines). Paris: Gallimard.
- VAN DEN BERG, B. (1949). "De conjugatie van het praesens in de Noordnederlandse dialecten". *Taal en Tongval* 1: 6-13.
- VAN DEN BERG, HELMA (1995). *A Grammar of Hunzib (with texts and lexicon)*. Ph.D. dissertation, Universiteit Leiden.
- (1999). "Gender and person agreement in Akusha Dargi". *Folia Linguistica* XXXIII (2): 153-168.
- BIGALKE, RAINER (1997). *Siciliano*. (Languages of the World/Materials, 129). München: Lincom.
- BIRK, D.B.W. (1976). *The Malakmalak Language: Daly River (Western Arnhem Land)*. (Pacific Linguistics, B 45). Canberra: Australian National University.
- BLAKE, BARRY J. (1988). "Redefining Pama-Nyungan: towards the prehistory of Australian Languages". *Aboriginal Linguistics* 1: 1-90.
- (1991). "The significance of pronouns in the history of Australian languages". In: Philip Baldi (ed.) *Patterns of Change, Change of Patterns*, pp. 219-234. Berlin: Mouton de Gruyter.
- BLAKE, FRANK R. (1934). "The Origin of Pronouns of the First and Second Person". *American Journal of Philology* 55 (219): 244-248.
- BLOOMFIELD, LEONARD (1942). "Outline of Ilocano Syntax". *Language* 18: 193-200.
- (1946). "Algonquian". In: Cornelius Osgood (ed.) *Linguistic Structures of Native America*, pp. 85-129. (Publications in Anthropology, 6). New York: Viking Fund.
- (1956). *Eastern Ojibwa: Grammatical sketch, texts and word list*. Ann Arbor:

- University of Michigan Press.
- (1962). *Menomini Language*. New Haven: Yale University Press.
- BOAS, FRANZ (1911a). “Chinook”. In: Franz Boas (ed.) *Handbook of American Indian Languages*. Vol. 1, pp. 559-678. Washington: Bureau of American Ethnology.
- (1911b). “Introduction”. In: Franz Boas (ed.) *Handbook of American Indian Languages*. Vol. 1, pp. 1-83. Washington: Bureau of American Ethnology.
- (1947). *Kwakiutl Grammar: with a glossary of the suffixes*. (Transactions of the American Philosophical Society, 37:3). Philadelphia: American Philosophical Society.
- BOHNHOFF, LEE E. (1986). “Yag Dii (Duru) Pronouns”. In: Ursula Wiesemann (ed.) *Pronominal Systems*, pp. 103-130. Tübingen: Gunter Narr.
- BOPP, FRANZ (1820). *Analytical Comparison of the Sanskrit, Greek, Latin and Teutonic Languages, showing the original identity of their grammatical structure*. (Amsterdam Classics in Linguistics, 3). Amsterdam: John Benjamins. (English translation by the author of his 1816 work in German, reprint 1974).
- BORGMAN, DONALD M. (1990). “Sanuma”. In: Desmond C. Debyshire & Geoffrey K. Pullum (eds.) *Handbook of Amazonian Languages*. Vol. 2, pp. 15-248. Berlin: Mouton de Gruyter.
- BOXWELL, MAURICE (1967). “Weri Pronoun System”. *Linguistics* 29: 34-43.
- DE BRAY, R.G.A. (1951). *Guide to the Slavonic Languages*. London: J.M. Dent & Sons.
- BROADBENT, SYLVIA M. (1964). *The Southern Sierra Miwok Language*. (University of California Publications in Linguistics, 38). Berkeley: University of California Press.
- BÜHLER, KARL (1934). *Sprachtheorie*. Jena: Fischer.
- BULATOVA, NADEZDA & LENORE GRENOBLE (1999). *Evenki*. (Languages of the World/Materials, 141). München: Lincom.
- BURLING, ROBBINS (1970). *Man's Many Voices*. New York: Holt, Rinehart & Winston.
- BURQUEST, DONALD A. (1986). “The Pronoun System of Some Chadic Languages”. In: Ursula Wiesemann (ed.) *Pronominal Systems*, pp. 71-102. Tübingen: Gunter Narr.
- BYBEE, JOAN L., WILLIAM PAGLIUCA & REVERE D. PERKINS (1990). “On the asymmetries in the affixation of grammatical material”. In: William Croft (ed.) *Studies in Typology and Diachrony*, pp. 1-42. (Typological Studies in Language, 20). Amsterdam: John Benjamins.
- CALAME-GRIAULE (1968). *Dictionnaire Dogon*. Paris: Librairie C. Klincksieck.
- CALDWELL, ROBERT (1856). *A Comparative Grammar of the Dravidian or South-Indian Family of Languages*. London: Kegan Paul, Trench, Trübner & Co. (reprint 1913).
- CALLAGHAN, CATHERINE A. (1965). *Lake Miwok Dictionary*. (University of California Publications in Linguistics, 39). Berkeley: University of California.
- (1974). “Increase in Morphological Complexity”. In: Luigi Heilmann (ed.) *Proceedings of the Eleventh International Congress of Linguistics*, pp. 383-388. Bologna: Società Editrice il Mulino.
- (1984). *Plains Miwok Dictionary*. (University of California Publications in Linguistics, 105). Berkeley: University of California Press.
- (1987). *Northern Sierra Miwok Dictionary*. (University of California Publications

- in Linguistics, 110). Berkeley: University of California Press.
- CALLOW, JOHN CAMPBELL (1962). *The Apinayé Language: Phonology and Grammar*. Ph.D. dissertation, University of London.
- CAMPBELL, LYLE (1985). *The Pipil Language of El Salvador*. (Mouton Grammar Library, 1). Berlin: Mouton.
- (1997). *American Indian Languages: The Historical Linguistics of Native America*. (Oxford Studies in Anthropological Linguistics, 4). Oxford: Oxford University Press.
- CAPELL, A. (1940). “Language Study for New Guinea Students”. *Oceania* 11 (1): 40-74.
- (1969a). “Non-Austronesian Languages of the British Solomons”. In: *Papers in Linguistics of Melanesia*. Vol. 2, pp. 1-16. (Pacific Linguistics, 21). Canberra: Australian National University.
- (1969b). “The Structure of the Binandere Verb”. In: *Papers in New Guinea Linguistics*. Vol. 9, pp. 1-32. (Pacific Linguistics, A 18). Canberra: Australian National University.
- (1975). “The ‘West Papuan Phylum’: General, and Timor and Areas Further West”. In: S.A. Wurm (ed.) *Papuan Languages and the New Guinea Linguistic Scene*, pp. 667-716. (Pacific Linguistics, C 38). Canberra: Australian National University.
- CARLIN, EITHNE B. (1997). “Verbal Qualities of Nouns and Nominalizations in Trio”. Paper presented at Amsterdam: ATW Colloquium (June 27, 1997).
- CHASE, WAYLAND JOHNSON (1926). *The Ars Minor of Donatus: For one thousand years the leading textbook of grammar, translated from the Latin with introductory sketch*. Madison: University of Wisconsin.
- CHELLIAH, SHOBHANA L. (1997). *A Grammar of Meithei*. (Mouton Grammar Library, 17). Berlin: Mouton de Gruyter.
- CHLENOVA, S.F. (1973). “Kategorija chisla v lichnyx mestoimenijax [Category of number within personal pronouns]”. In: *Lingvotipologicheskie issledovanija*, pp. 164-201. (Vypusk 1, Chast' 1). Moskva: Izdatel'stvo MGU.
- CHOMSKY, NOAM (1981). *Lectures on Government and Binding*. (Studies in Generative Grammar). Dordrecht: Foris.
- COLE, PETER (1982). *Imbabura Quechua*. (Lingua Descriptive Studies, 5). Amsterdam: North Holland Publishing Company.
- COMRIE, BERNARD (1980a). “Inverse Verb Forms in Siberia: evidence from Chukchee, Koryak and Kamchadal”. *Folia Linguistica Historica* 1 (1): 61-74.
- (1980b). “review of: Greenberg, Joseph H. (ed.) *Universals of Human Language*. Vol. 3: word structure”. *Language* 56 (4): 834-838.
- (1989). *Language Universals and Linguistic Typology*. Oxford: Blackwell. (2nd edition).
- COMRIE, BERNARD, MARIA POLINSKY & RAMAZAN RAJABOV (1999). “Tsezian Languages”. Unpublished manuscript January 11, 1999. (reader distributed at LOT winterschool).
- CONKLIN, HAROLD C. (1962). “Lexicographical Treatment of Folk Taxonomies”. In: Fred W. Householder & Sol Saporta (eds.) *Problems in Lexicography*, pp. 119-142. Bloomington: Indiana University Press. (third edition, 1975).
- COOK, WILLIAM HINTON (1979). *A Grammar of North Carolina Cherokee*. Ph.D. dissertation, University of Ann Arbor, Michigan.
- COOKE, JOSEPH R. (1968). *Pronominal Reference in Thai, Burmese, and Vietnamese*.

- (University of California Publications in Linguistics, 52). Berkeley: University of California Press.
- CORBETT, GREVILLE G. (1979). "The Agreement Hierarchy". *Journal of Linguistics* 15: 203-224.
- (1991). *Gender*. (Cambridge Textbooks in Linguistics). Cambridge: Cambridge University Press.
- (1994). "Agreement". In: R.E. Asher (ed.) *The Encyclopedia of Language and Linguistics*, pp. 54-60. Oxford: Pergamon Press.
- (1997). "A typology of nominal number systems: values and constraints". Paper presented at Eugene, Oregon: ALT II (September 12, 1997).
- CORBETT, GREVILLE G. & MARIANNE MITHUN (1996). "Associative forms in a typology of number systems: evidence from Yup'ik". *Journal of Linguistics* 32: 1-17.
- COWAN, H.K.J. (1965). *Grammar of the Sentani Language*. (Verhandelingen van het Koninklijk Instituut voor Taal-, Land- en Volkenkunde, 47). 's Gravenhage: Martinus Nijhoff.
- CRAZZOLARA, J.P. (1955). *A Study of the Acooli Language*. London: Oxford University Press.
- (1960). *A Study of the Logbara (Ma'di) Language*. London: Oxford University Press.
- CROFT, WILLIAM (1990). *Typology and Universals*. (Cambridge Textbooks in Linguistics). Cambridge: Cambridge University Press.
- CROFTS, MARJORIE (1973). *Gramática Mundurukú*. (Séria Linguística, 2). Brasília: Summer Institute of Linguistics.
- CROWELL, THOMAS HARRIS (1979). *A Grammar of Bororo*. Ph.D. dissertation, Cornell University.
- CROWLEY, TERRY (1981). "The Mpakwithi dialect of Anguthimri". In: R.M.W. Dixon & Barry J. Blake (eds.) *Handbook of Australian Languages*. Vol. 2, pp. 147-194. Amsterdam: John Benjamins.
- (1982). *The Paamese Language of Vanuatu*. (Pacific Linguistics, B 87). Canberra: Australian National University.
- (1983). "Uradhi". In: R.M.W. Dixon & Barry J. Blake (eds.) *Handbook of Australian Languages*. Vol. 3, pp. 307-430. Amsterdam: John Benjamins.
- CSÚCS, SÁNDOR (1998). "Udmurts". In: Daniel Abondolo (ed.) *The Uralic Languages*, pp. 276-304. London: Routledge.
- CULICOVER, PETER W. (ed.) (1997). *Principles and Parameters, an Introduction to Syntactic Theory*. Oxford: Oxford University Press.
- D'ALTON, PAULA AGNES (1983). *Esquisse Phonologique et Grammaticale du Palor, Langue Cangin (Senegal)*. Ph.D. dissertation, Université de la Sorbonne Nouvelle, Paris.
- DELANCEY, SCOTT (1987). "Sino-Tibetan Languages". In: Bernard Comrie (ed.) *The World's Major Languages*, pp. 797-810. London: Croom Helm.
- DERBYSHIRE, DESMOND C. (1979). *Hixkaryana*. Amsterdam: North Holland Publishing Company.
- DIBBETS, G.R.W. (1995). *De Woordsoorten in de Nederlandse Triviumgrammatica*. (Stichting Neerlandistiek VU, 18). Münster: Nodus.
- DIMMENDAAL, GERRIT J. (1982). *The Turkana Language*. Ph.D. dissertation, Universiteit Leiden.
- (1983). "Topics in a Grammar of Turkana". In: M. Lionel Bender (ed.) *Nilo-*

- Saharan Language Studies*, pp. 239-271. Michigan: African Studies Center.
- DIXON, R.M.W. (1972). *The Dyirbal Language of North Queensland*. Cambridge: Cambridge University Press.
- (1977). *A Grammar of Yidjñ*. Cambridge: Cambridge University Press.
- (1979). “Ergativity”. *Language* 55: 59-138.
- (1980). *The Languages of Australia*. Cambridge: Cambridge University Press.
- (1981). “Wargamay”. In: R.M.W. Dixon & Barry J. Blake (eds.) *Handbook of Australian Languages*. Vol. 2, pp. 1-144. Amsterdam: John Benjamins.
- (1983). “Nyawaygi”. In: R.M.W. Dixon & Barry J. Blake (eds.) *Handbook of Australian Languages*. Vol. 3, pp. 431-525. Amsterdam: John Benjamins.
- (1997). *The Rise and Fall of Languages*. Cambridge: Cambridge University Press.
- DIXON, R.M.W. & BARRY J. BLAKE (1979). “Introduction”. In: R.M.W. Dixon & Barry J. Blake (eds.) *Handbook of Australian Languages*. Vol. 1, pp. 1-26. Amsterdam: John Benjamins.
- DONAHUE, MARK (1999). *Warembori*. (Languages of the World/Materials, 341). München: Lincom.
- DONEUX, J.L. (1968). *Esquisse Grammaticale de Dan*. (Documents Linguistiques, 15). Dakar: Université de Dakar.
- DORSCH, H. (1911). “Grammatik der Nkosi-Sprache”. *Zeitschrift für Kolonialsprachen* 1: 241-283.
- VAN DRIEM, GEORGE (1987). *A Grammar of Limbu*. (Mouton Grammar Library, 4). Berlin: Mouton de Gruyter.
- DU FEU, VERONICA (1996). *Rapanui*. London: Routledge.
- DURIE, MARK (1985). *A Grammar of Acehnese: on the basis of a dialect of North Aceh*. (Verhandelingen van het Koninklijk Instituut voor Taal-, Land- en Volkenkunde, 112). Dordrecht: Foris.
- DUTTON, TOM E. (1996). *Koiari*. (Languages of the World/Materials, 10). München: Lincom.
- EBERMANN, E. (1986). *Die Sprache der Mauka*. Ph.D. dissertation, University of Vienna.
- EBERT, KAREN (1997). *Camling (Chamling)*. (Languages of the World/Materials, 103). München: Lincom.
- EISENBERG, PETER (1994). “German”. In: Ekkehard König & Johan van der Auwera (eds.) *The Germanic Languages*, pp. 349-387. London: Routledge.
- ELDERKIN, EDWARD D. (1988). “Person and number markers in Iragw verbs”. *Afrikanische Arbeitspapiere* 14: 79-96.
- ELSON, BEN (1960). “Sierra Popoluca Morphology”. *International Journal of American Linguistics* 26 (3): 206-223.
- EMENEAU, MURRAY B. (1984). *Toda grammar and texts*. Philadelphia: American Philosophical Society.
- ESSER, S.J. (1964). *De Uma-taal (West Midden-Celebes): Spraakkunstige schets en teksten*. (Verhandelingen van het Koninklijk Instituut voor Taal-, Land- en Volkenkunde, 43). 's-Gravenhage: Martinus Nijhoff.
- EVERETT, DANIEL L. (1986). “Pirahã”. In: Desmond C. Debyshire & Geoffrey K. Pullum (eds.) *Handbook of Amazonian Languages*. Vol. 1, pp. 200-325. Berlin: Mouton de Gruyter.
- (1987). “Pirahã Clitic Doubling”. *Natural Language and Linguistic Theory* 5: 245-276.
- EVERETT, DANIEL L. & SARAH G. THOMASON (n.d.). “Borrowed pronouns in Piraha”.

- Unpublished manuscript. (Paper available at <http://amazonling.linguist.pitt.edu/pirdiac.html>).
- FARR, JAMES & CYNTHIA FARR (1975). "Some Features of Korafe Morphology". In: Tom E. Dutton (ed.) *Studies in Language of Central and South-East Papua*, pp. 731-770. (Pacific Linguistics, C 29). Canberra: Australian National University.
- FOLEY, WILLIAM A. (1986). *The Papuan Languages of New Guinea*. (Cambridge Language Surveys). Cambridge: Cambridge University Press.
- (1991). *The Yimas Language of New Guinea*. Stanford: Stanford University Press.
- (1997). *Anthropological Linguistics: An Introduction*. (Language in Society, 24). Oxford: Blackwell.
- FORCHHEIMER, PAUL (1953). *The Category of Person in Language*. Berlin: Walter de Gruyter.
- FORTESCUE, MICHAEL (1984). *West Greenlandic*. (Croom Helm Descriptive Grammars). London: Croom Helm.
- FORTUNE, R.F. (1942). *Arapesh*. New York: J.J. Augustin.
- FOSTER, MARY L. & GEORGE M. FOSTER (1948). *Sierra Popoluca Speech*. (Smithsonian Institution Publications, 8). Washington D.C.: U.S. Government Printing Office.
- FOX, G.J. (1976). *Big Nambas Grammar*. (Pacific Linguistics, B 60). Canberra: Australian National University.
- FRACHTENBERG, LEO J. (1922a). "Coos". In: Franz Boas (ed.) *Handbook of American Indian Languages*. Vol. 2, pp. 297-430. Washington: Bureau of American Ethnology.
- (1922b). "Siuslawan (Lower Umpqua)". In: Franz Boas (ed.) *Handbook of American Indian Languages*. Vol. 2, pp. 441-629. Washington: Bureau of American Ethnology.
- FRANK, PAUL S. (1985). *A Grammar of Ika*. Ph.D. dissertation, University of Pennsylvania.
- (1990). *Ika Syntax: Studies in the languages of Colombia*. Arlington: Summer Institute of Linguistics.
- FRANKLIN, K.J. (1971). *A Grammar of Kewa, New Guinea*. (Pacific Linguistics, C 16). Canberra: Australian National University.
- FRANSEN, MARGO A.E. (1995). *A Grammar of Limbum, a Grassfields Bantu Language spoken in the North-West Province of Cameroon*. Ph.D. dissertation, Vrije Universiteit Amsterdam.
- FRANTZ, CHESTER & HOWARD P. MCKAUGHAN (1973). "Gadsup Independent Verb Affixes". In: Howard P. McKaughan (ed.) *The Languages of the Eastern Family of the East New Guinea Highland Stock*, pp. 439-449. Seattle: University of Washington Press.
- FREELAND, L.S. (1947). "Western Miwok texts with linguistic sketch". *International Journal of American Linguistics* 13: 31-46.
- (1951). *Language of the Sierra Miwok*. (Memoir of the International Journal of American Linguistics, 6). Baltimore: Waverly Press.
- GADEN, H. (1909). *Essai de Grammaire de la Langue Baguirmienne*. Paris: Ernest Leroux.
- GAIR, JAMES W. (1970). *Colloquial Sinhalese Clause Structures*. (Janua Linguarum, Series Practica, 83). The Hague: Mouton.
- GARCIA, ERICA C. & FLORIMON VAN PUTTE (1989). "Forms are Silver, Nothing is

- Gold". *Folia Linguistica Historica* 8 (1-2): 365-384.
- GARDINER, ALAN (1932). *The Theory of Speech and Language*. Oxford: Clarendon Press. (second edition 1951).
- GARLAND, ROGER & SUSAN GARLAND (1975). "A Grammar Sketch of Moutain Koiali". In: Tom E. Dutton (ed.) *Studies in Language of Central and South-East Papua*, pp. 413-470. (Pacific Linguistics, C 29). Canberra: Australian National University.
- GARY, J.O. & S. GAMAL-ELDIN (1982). *Cairene Egyptian Colloquial Arabic*. (Lingua descriptive Studies). Amsterdam: North Holland Publishing Company.
- GEARY, ELAINE (1977). *Kunimaipa Grammar*. (Workpapers in Papua New Guinean Languages, 23). Ukarumpa: Summer Institute of Linguistics.
- GEOGHEGAN, RICHARD HENRY (1944). *The Aleut Language*. Washington: United States Department of the Interior.
- GEORGE, KEN (1993). "Cornish". In: Martin J. Ball & Fife James (eds.) *The Celtic Languages*, pp. 410-468. London: Routledge.
- GERZENSTEIN, ANA (1994). *Lengua Maká: Estudio descriptivo*. (Archivo de Lenguas Indoamericanas). Buenos Aires: Universidad de Buenos Aires.
- GIVÓN, T. (1976). "Topic, Pronoun and Grammatical Agreement". In: Charles N. Li (ed.) *Subject and Topic*, pp. 149-188. New York: Academic Press.
- (1980). *Ute Reference Grammar*. Ignacio: Ute Press.
- (1984). *Syntax, a Functional-Typological Introduction*. Vol. 1. Amsterdam: John Benjamins.
- (1990). *Syntax, a Functional-Typological Introduction*. Vol. 2. Amsterdam: John Benjamins.
- GLASGOW, KATHLEEN (1964). "Four principal contrasts in Burera Personal Pronouns". In: Richard Pittman & Harland B. Kerr (eds.) *Papers on the Languages of the Australian Aborigines*, pp. 109-117. (Occasional Papers in Aboriginal Studies, 3). Canberra: Australian Institute of Aboriginal Studies.
- (1984). "Burarra word classes". In: *Papers in Australian Linguistics*. Vol. 16, pp. 1-54. (Pacific Linguistics, A 68). Canberra: Australian National University.
- GODDARD, PLINY EARLE (1905). *The Morphology of the Hupa language*. (American Archeology and Ethnology). Berkeley: University Press.
- (1912). *Elements of the Kato language*. Berkeley: University of California Press.
- GODFREY, MARIE (1964). "A Tentative Outline Grammar of Wik-Munkan". In: *Occasional Papers in Aboriginal Studies*. Vol. 2, pp. 57-78. Canberra: Australian Institute of Aboriginal Studies.
- GODFREY, MARIE & HARLAND B. KERR (1964). "Personal Pronouns in Wik-Munkan". In: Richard Pittman & Harland B. Kerr (eds.) *Papers on the Languages of the Australian Aborigines*, pp. 13-34. (Occasional Papers in Aboriginal Studies, 3). Canberra: Australian Institute of Aboriginal Studies.
- GOFFMAN, ERVING (1979). "Footing". *Semiotica* 25 (1/2): 1-29.
- GONZALES, ANDREW (1981). *Pampangan: towards a meaning-based description*. (Pacific Linguistics, C 48). Canberra: Australian National University.
- GONZÁLES DE PÉREZ, MARÍA STELLA (1987). *Diccionario y Gramática Chibcha: manuscrito anónimo de la bibliotheca nacional de Colombia*. (Bibliotheca 'Ezequiel Uricoechea', 1). Bogotá: Instituto Caro Y Cuervo.
- GORDON, LYNN (1986). *Maricopa Morphology and Syntax*. Berkeley: University of California Press.
- GÖRLACH, MANFRED (1997). *The Linguist History of English: An Introduction*.

- London: Macmillian.
- GRALOW, FRANCES L. (1993). *Un Bosquejo del Idioma Koreguaje*. Santafé de Bogotá: Instituto lingüístico de Verano.
- GREENBERG, JOSEPH H. (1963). "Some universals of grammar with particular reference to the order of meaningful elements". In: Joseph H. Greenberg (ed.) *Universals of Language*, pp. 73-113. Cambridge, Mass.: MIT Press.
- (1966). *Language Universals: with special reference to feature hierarchies*. (Janua Linguarum, series minor, LIX). The Hague: Mouton.
- (1969). "Some Methods of Dynamic Comparison in Linguistics". In: Jaan Puhvel (ed.) *Substance and Structure of Language*, pp. 147-204. Berkeley: University of California Press.
- (1985). "Some Iconic Relationship among Place, Time, and Discourse Deixis". In: John Haiman (ed.) *Iconicity in Syntax*, pp. 271-288. (Typological Studies in Language, 6). Amsterdam: John Benjamins.
- (1988). "The first person inclusive dual as an ambiguous category". *Studies in Language* 12 (1): 1-18.
- (1989). "On a metalanguage for pronominal systems: a reply to McGregor". *Studies in Language* 13 (2): 452-458.
- (1993). "The second person is rightly so called". In: Mushira Eid & Gregory Iverson (eds.) *Principles and Prediction: the analysis of natural language. Papers in the honor of Gerald Sanders*, pp. 9-24. Amsterdam: John Benjamins.
- GRIMES, BARBARA F. (ed.) (1996). *Ethnologue*. Dallas: Summer Institute of Linguistics. (13).
- DE GROOT, CASPER & MACHIEL J. LIMBURG (1986). *Pronominal elements: diachrony, typology, and formalization in Functional Grammar*. (Working Papers in Functional Grammar, 12). Amsterdam: IFOTT.
- GRUZDEVA, EKATERINA (1998). *Nivkh*. (Languages of the World/Materials, 111). München: Lincom.
- HAACKE, W.H.G. (1977). "The so-called 'personal pronoun' in Nama". In: A. Traill (ed.) *Khoisan Linguistic Studies*. Vol. 3, pp. 43-62. Johannesburg: African Studies Institute.
- HAAS, MARY R. (1946). "A Grammatical Sketch of Tunica". In: Cornelius Osgood (ed.) *Linguistic Structures of Native America*, pp. 337-366. (Publications in Anthropology, 6). New York: Viking Fund.
- (1969). "'Exclusive' and 'Inclusive': a look at early usage". *International Journal of American Linguistics* 35 (1): 1-6.
- (1977). "From Auxiliary Verb Phrase to Inflectional Suffix". In: Charles N. Li (ed.) *Mechanisms of Syntactic Change*, pp. 525-538. Austin: University of Texas Press.
- HAGÈGE, CLAUDE (1982). *La Structure des Langues*. (Que sais-je?, 2006). Paris: Presses Universitaires de France.
- HAGMAN, ROY S. (1977). *Nama Hottentot Grammar*. Bloomington: Indiana University Press.
- HAIMAN, JOHN (1980). *Hua: A Papuan Language of the Eastern Highlands of New Guinea*. (Studies in Language Companion Series, 5). Amsterdam: John Benjamins.
- HALE, KENNETH (1973). "Person Marking in Walbiri". In: Paul Kiparski (ed.) *A Festschrift for Morris Halle*, pp. 308-344. New York: Holt, Rinehart &

- Winston.
- (1976a). “Tya.pukay (Djaabugay)”. In: Peter Sutton (ed.) *Languages of Cape York*, pp. 236-242. Canberra: Australian Institute of Aboriginal Studies.
- (1976b). “Wik reflections of Middle Paman Phonology”. In: Peter Sutton (ed.) *Languages of Cape York*, pp. 50-60. Canberra: Australian Institute of Aboriginal Studies.
- HALLEBEEK, J., A. VAN BOMMEL & C. VAN ESCH (1994). *Moderne Spaanse grammatica*. Groningen: Wolters-Noordhoff.
- HALPERN, A.M. (1946). “Yuma”. In: Cornelius Osgood (ed.) *Linguistic structures of Native America*, pp. 249-288. (Publications in Anthropology, 6). New York: Viking Fund.
- HARBERT, WAYNE (1995). “Binding Theory, Control, and pro”. In: Gert Webelhuth (ed.) *Government and Binding Theory and the Minimalist Program*, pp. 177-240. Oxford: Blackwell.
- HARDMAN, M.J. (1966). *Jaqaru: outline of phonological and morphological structure*. (Janua Linguarum, Series Practica, 22). The Hague: Mouton.
- (1972). “Early Use of Inclusive/Exclusive”. *International Journal of American Linguistics* 38 (2): 145-146.
- HARLOW, RAY (1996). *Maori*. (Languages of the World/Materials, 20). München: Lincom.
- HARMS, PHILLIP LEE (1994). *Epena Pedee Syntax*. (Studies in the Languages of Colombia, 4). Arlington: Summer Institute of Linguistics.
- HARRISON, CARL H. (1971). “The morphophonology of Asurini words”. In: David Bendor-Samuel (ed.) *Tupi Studies*. Vol. 1. (Summer Institute of Linguistics Publications in Linguistics and Related Fields, 29). Oklahoma: Summer Institute of Linguistics.
- HASPELMATH, MARTIN (1993). *A Grammar of Lezgian*. (Mouton Grammar Library, 9). Berlin: Mouton de Gruyter.
- (1997). *Indefinite Pronouns*. (Oxford Studies in Typology and Linguistic Theory). Oxford: Clarendon Press.
- HAVILAND, JOHN (1979). “Guugu-Yimidhirr”. In: R.M.W. Dixon & Barry J. Blake (eds.) *Handbook of Australian Languages*. Vol. 1, pp. 27-180. Amsterdam: John Benjamins.
- HAWKINS, ROBERT E. (1998). “Wai Wai”. In: Desmond C. Derbyshire & Geoffrey K. Pullum (eds.) *Handbook of Amazonian Languages*. Vol. 4, pp. 25-224. Berlin: Mouton de Gruyter.
- HEAD, BRIAN F. (1978). “Respect Degrees in Pronominal Reference”. In: Joseph H. Greenberg (ed.) *Universals of Human Language*. Vol. 3: word structure, pp. 151-212. Stanford: Stanford University Press.
- HEALEY, ALAN, AMBROSE ISOROEMBO & MARTIN CHITTLEBOROUGH (1969). “Preliminary Notes on Orokaiva Grammar”. In: *Papers in New Guinea Linguistics*. Vol. 9, pp. 33-64. (Pacific Linguistics, A 18). Canberra: Australian National University.
- HEATH, JEFFREY (1978). *Ngandi grammar, texts, and dictionary*. Canberra: Australian Institute of Aboriginal Studies.
- (1984). *Functional Grammar of Nunggbuyu*. Canberra: Australian Institute of Aboriginal Studies.
- (1991). “Pragmatic disguise in pronominal-affix paradigms”. In: Frans Plank (ed.) *Paradigms, The Economy of Inflection*, pp. 75-89. Berlin: Mouton de Gruyter.

- (1998). “Pragmatic Skewing in 1 ↔ 2 pronominal combination in native American languages”. *International Journal of American Linguistics* 64 (2): 83-104.
- HELMISKI, EUGENE (1998). “Nganasan”. In: Daniel Abondolo (ed.) *The Uralic Languages*, pp. 480-515. London: Routledge.
- HELMBRECHT, JOHANNES (1996). “The syntax of personal agreement in East Caucasian languages”. *Sprachtypologie und Universalien Forschung* 49 (2): 127-148.
- (1999). “The typology of 1st person marking and its cognitive backgrounds”. In: Masoko K. Hiraga, Chris Sinha & Sherman Wilcox (eds.) *Cultural, Psychological and Typological Issues in Cognitive Linguistics*, pp. 285-297. (Current Issues in Linguistic Theory, 152). Amsterdam: John Benjamins.
- HENDERSON, T.S.T. (1985). “Who are *we*, anyway? A study of personal pronoun systems”. *Linguistische Berichte* 98: 300-309.
- HENGEVELD, KEES (1997). “Shifters”. Paper presented at Inaugurele rede Universiteit van Amsterdam (31 october 1997).
- HEWITT, B. GEORGE (1979). *Abkhaz*. (Lingua Descriptive Studies). Amsterdam: North Holland Publishing Company.
- (ed.) (1989). *The North West Caucasian Languages*. (The Indigenous Languages of the Caucasus, 2). Delmar: Caravan Books.
- HODDINOTT, W.G. & F.M. KOFOD (1988). *The Ngankikurungkurr Language*. (Pacific Linguistics, D 77). Canberra: Australian National University.
- HOFF, B.J. (1968). *The Carib Language*. (Verhandelingen van het Koninklijk Instituut voor Taal-, Land- en Volkenkunde, 55). Den Haag: Martinus Nijhoff.
- HOFFMANN, C. (1963). *A Grammar of the Margi Language*. London: Oxford University Press.
- HOFFMANN, J. (1903). *Mundari Grammar*. Calcutta: Bengal Secretariat Press.
- HOJER, HARRY (1946). “Chiricahua Apache”. In: Cornelius Osgood (ed.) *Linguistic Structures of Native America*. (Publications in Anthropology, 6). New York: Viking Fund.
- HOLLENBACH, BARBARA ERICKSON (1970). “Inclusive Plural: A Further Look”. *Linguistics* 60: 27-32.
- HOLMER, NILS M. (1946). “Outline of Cuna Grammar”. *International Journal of American Linguistics* 12 (4): 185-197.
- HOPKINS, ELIZABETH B. (1986). “Pronouns and Pronoun Fusion in Yaouré”. In: Ursula Wiesemann (ed.) *Pronominal Systems*, pp. 191-204. Tübingen: Gunter Narr.
- HOSKISON, J.T. (1983). *A Grammar and Dictionary of the Gude Language*. Ph.D. dissertation, Ohio State University.
- HOUSEHOLDER, FRED W. (1955). “review of: Paul Forchheimer *The Category of Person in Language*”. *Language* 31 (1-2): 93-100.
- HOWE, STEPHEN (1996). *The Personal Pronouns in the Germanic Languages: A study of personal pronoun morphology and change in the Germanic languages from the first records to the present day*. (Studia Linguistica Germanica, 43). Berlin: Walter de Gruyter.
- HOYER, HARRY (1933). *Tonkawa, an Indian language of Texas*. New York: Columbia University Press.
- HUANG, YAN (1995). “On Null Subject and Null Objects in Generative Grammar”. *Linguistics* 33: 1081-1123.

- VON HUMBOLDT, WILHELM (1827). "Über den Dualis". In: *Über die Sprache, Reden vor der Akademie*, pp. 143-169. Tübingen: Francke. (1994).
- (1830). *Über die Verwandtschaft der Ortsadverbien mit dem Pronomen in einigen Sprachen*. Berlin: Königliche Akademie der Wissenschaften.
- HUTCHISSON, DON (1986). "Sursurunga Pronouns and the Special Uses of Quadral number". In: Ursula Wieseemann (ed.) *Pronominal Systems*, pp. 1-20. Tübingen: Gunter Narr.
- HYMAN, LARRY M. (1979). "Phonology and Noun Structure". In: Larry M. Hyman (ed.) *Aghem Grammatical Structure*, pp. 1-72. (Southern California Occasional Papers in Linguistics, 7). Los Angeles: University of Southern California.
- (1981). *Noni Grammatical Structure: with special reference to verb morphology*. (Southern California Occasional Papers in Linguistics, 9). Los Angeles: University of Southern California.
- HYMES, DELL H. (1955). "review of: Paul Forchheimer *The Category of Person in Language*". *International Journal of American Linguistics* 21 (3): 294-300.
- (1972). "On Personal Pronouns: 'Fourth' Person and Phonesthematic Aspects". In: E. Estellie Smith (ed.) *Studies in Linguistics in Honor of George L. Trager*, pp. 100-121. (Janua Linguarum, Series Maior, 52). The Hague: Mouton.
- INGRAM, DAVID (1978). "Typology and Universals of Personal Pronouns". In: Joseph H. Greenberg (ed.) *Universals of Human Language*. Vol. 3: word structure, pp. 213-248. Stanford: Stanford University Press.
- INNES, G. (1966). *An Introduction to Grebo*. London: School of Oriental and African Studies.
- IRWIN, BARRY (1974). *Salt-Yui Grammar*. (Pacific Linguistics, B 35). Canberra: Australian National University.
- JAKOBSON, ROMAN (1971). "Shifters, verbal categories and the Russian verb". In: Roman Jakobson (ed.) *Selected Writings*. Vol. 2, pp. 130-147. Den Haag: Mouton.
- JELINEK, ELOISE (1984). "Empty Categories, Case, and Configurationality". *Natural Language and Linguistic Theory* 2 (1): 39-76.
- JENSEN, CHERYL (1990). "Cross-Referencing Changes in some Tupí-Guaraní Languages". In: Doris L. Payne (ed.) *Amazonian Linguistics*, pp. 117-160. Austin: University of Texas Press.
- JESPERSEN, OTTO (1922). *Language, it's Nature, Development and Origin*. London: George Allen & Unwin LTD.
- (1924). *The Philosophy of Grammar*. London: George Allen & Unwin LTD.
- JONES, WENDELL & PAULA JONES (1991). *Barasano Syntax*. (Studies in the Languages of Colombia). Arlington: Summer Institute of Linguistics.
- KANSTENKOLZ, R. (1987). *Das Koranko*. Ph.D. dissertation, Universität zu Köln.
- KESING, ROGER M. (1985). *Kwaio Grammar*. (Pacific Linguistics, B 88). Canberra: Australian National University.
- KEMP, ALAN (1987). "The Tehknè Grammatikè of Dionysius Thrax, translated into English". In: Daniel J. Taylor (ed.) *The History of Linguistics in the Classical Period*, pp. 169-189. (Studies in the History of the Language Sciences, 46). Amsterdam: John Benjamins.
- KENDALL, MARTHA B. (1976). *Selected Problems in Yavapai Syntax*. New York: Garland.
- KERESZTES, LÁSZLÓ (1998). "Mansi". In: Daniel Abondolo (ed.) *The Uralic*

- Languages*, pp. 387-427. London: Routledge.
- VAN DE KERKE, SIMON (1996). *Affix Order and Interpretation in Bolivian Quechua*. Ph.D. dissertation, Universiteit van Amsterdam.
- KIEBLING, ROLAND (1994). *Eine Grammatik des Burunge*. (Afrikanistische Forschungen). Hamburg: Reasearch and Progress Verlag.
- KIMBALL, GEOFFREY DAVID (1985). *A Descriptive grammar of Koasati*. Ph.D. dissertation, Tulane University.
- KIRK, J.W.C. (1905). *A grammar of the Somali language*. Cambridge: Cambridge University Press.
- KITCHING, A.L. (1907). *An outline Grammar of the Acholi Language*. London: The Sheldon press.
- KOCH, HARALD (1994). "The Creation of Morphological Zeroes". In: Geert Booij & Jaap van Marle (eds.) *Yearbook of Morphology 1994*, pp. 31-71. Dordrecht: Kluwer.
- (1997). "Morphological changes in Arandic pronouns". Paper presented at Düsseldorf: the XIII International Conference on Historical Linguistics (August 15, 1997).
- KOEHN, EDWARD & SALLY KOEHN (1986). "Apalai". In: Desmond C. Debyshire & Geoffrey K. Pullum (eds.) *Handbook of Amazonian Languages*. Vol. 1, pp. 33-127. Berlin: Mouton de Gruyter.
- KUHN, THOMAS S. (1962). *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press.
- KUIPERS, AERT H. (1974). *The Shuswap Language: Grammar, Texts, Dictionary*. The Hague: Mouton.
- KURYLOWICZ, JERZY (1964). *The inflectional categories of Indo-European*. Heidelberg: Carl Winter.
- LANGDON, MARGARET (1970). *A Grammar of Diegueño*. Berkeley: University of California Press.
- LAYCOCK, DON (1977). "Me and you versus the rest". *Irian* 6: 33-41.
- LEHMANN, CHRISTIAN (1982). "Universal and Typological Aspects of Agreement". In: Hansjakob Seiler & Franz Jozef Stachowiak (eds.) *Apprehension. Das Sprachliche Erfassen von Gegenständen*. Vol. 2: Die Techniken und ihr Zusammenhang in Einzalsprachen, pp. 201-267. Gunter Narr: Tübingen.
- (1983). "Rektion und Syntaktische Relationen". *Folia Linguistica* XVII: 339-378.
- LEVIN, NORMAN B. (1964). *The Assiniboine Language*. The Hague: Mouton.
- LEVINSON, STEPHEN C. (1988). "Putting Linguistics on a Proper Footing: Explorations in Goffman's Concepts of Participation". In: Paul Drew & Anthony Wootton (eds.) *Erving Goffman: Exploring the Interaction Order*, pp. 161-227. Oxford: Polity Press.
- LEWIS, G.L. (1967). *Turkish Grammar*. Oxford: Clarendon Press.
- LI, CHARLES N. & SANDRA A. THOMPSON (1979). "Third-person Pronouns and Zero-anaphora in Chinese Discourse". In: T. Givón (ed.) *Discourse and Syntax*, pp. 311-336. (Syntax and Semantics, 12). New York: Academic Press.
- LINDSTROM, LAMONT & JOHN D. LYNCH (1994). *Kwamera*. (Languages of the World/Materials, 2). München: Lincom.
- LOUNSBURY, FLOYD G (1953). *Oneida Verb Morphology*. (Yale University Publications in Anthropology, 48). New Haven: Yale University Press.
- LOVING, RICHARD (1973). "An Outline of Awa grammatical structure". In: Howard P. McKaughan (ed.) *The Languages of the Eastern Family of the East New*

- Guinea Highland Stock*, pp. 65-87. Seattle: University of Washington Press.
- LOWIE, ROBERT H. (1941). *The Crow Language: Grammatical Sketch and Analyzed Text*. Berkeley: University of California Press.
- LUHRMAN, G.J. (1984). *C.L. Papius, T. Linaker, J.C. Scaliger en hun beschouwing van het werkwoord. Een kritisch-vergelijkende studie omtrent XVIde eeuwse taalkundige theorievorming*. Ph.D. dissertation, Universiteit Groningen.
- LUKAS, J. (1953). *Die Sprache der Tubu in der zentralen Sahara*. Berlin: Akademie Verlag.
- LUPARDUS, KAREN JACQUE (1982). *The Language of the Alabama Indians*. Ph.D. dissertation, University of Kansas.
- LUSTED, M. (1976). "Anywa". In: Lionel M. Bender (ed.) *The Non-Semitic Languages of Ethiopia*, pp. 495-512. East Lansing: African Studies Center.
- LYNCH, JOHN D. (1967). *A Comparative Study of the Languages and Dialects of the Island of Tanna, Southern New Hebrides*. Ph.D. dissertation, University of Sydney.
- (1978). *A Grammar of Lenakel*. (Pacific Linguistics, B 55). Canberra: Australian National University.
- (1998). *Pacific Languages: An Introduction*. Honolulu: University of Hawai'i Press.
- LYONS, JOHN (1968). *Introduction to Theoretical Linguistics*. Cambridge: Cambridge University Press.
- (1977). *Semantics*. Cambridge: Cambridge University Press.
- MACLEAN, EDNA A. (1986). *North Slope Iñupiaq Grammar*. Fairbanks: Alaska Native Language Center. (3rd edition).
- MACUCH, RUDOLF (1965). *Handbook of Classical and Modern Mandaic*. Berlin: Walter de Gruyter.
- MALCHUKOV, ANDREI L. (1995). *Even*. (Languages of the World/Materials, 12). München: Lincom.
- MALONE, TERRELL (1988). "The Origin and Development of Tuyuca Evidentials". *International Journal of American Linguistics* 54 (2): 119-140.
- MANNHEIM, BRUCE (1982). "Person, Number and Inclusivity in Two Andean Languages". *Acta linguistica Hafniensia* 17 (2): 139-156.
- MANRIQUE, C. LEONARDO (1967). "Jiliapan Pame". In: Robert Wauchope & Norman A. McQuown (eds.) *Handbook of Middle American Indians*. Vol. 5: Linguistics, pp. 331-348. Austin: University of Texas Press.
- MARCHESE, LYNELL (1978). *Atlas Linguistique Kru: Essai de Typologie*. Abidjan: Université d'Abidjan.
- MARING, JOEL MARVYL (1967). *Grammar of Acoma Keresan*. Ph.D. dissertation, Indiana University.
- MASICA, COLIN P. (1991). *The Indo-Aryan Languages*. (Cambridge Language Surveys). Cambridge: Cambridge University Press.
- MATTHEWS, G.H. (1965). *Hidatsa Syntax*. (Papers in formal linguistics, 3). The Hague: Mouton.
- MCCELHANON, K.A. (1975). "North Eastern Trans-New Guinea Phylum Languages". In: S.A. Wurm (ed.) *Papuan Languages and the New Guinea Linguistic Scene*, pp. 527-568. (Pacific Linguistics, C 38). Canberra: Australian National University.
- MCGREGOR, WILLIAM B. (1989). "Greenberg on the first person inclusive dual: Evidence from some Australian languages". *Studies in Language* 13 (2): 437-

- 451.
- (1990). *A Functional Grammar of Gooniyandi*. (Studies in Language Companion Series, 22). Amsterdam: John Benjamins.
- (1993). *Gunin/Kwini*. (Languages of the World/Materials, 11). München: Lincom.
- (1994). *Warrwa*. (Languages of the World/Materials, 89). München: Lincom.
- (1996). *Nyulnyul*. (Languages of the World/Materials, 88). München: Lincom.
- MCKAUGHAN, HOWARD P. (1959). “Semantic Components of Pronoun Systems: Maranao”. *Word* 15: 101-102.
- MCKAY, GRAHAM R. (1978). “Pronominal person and number categories in Rembarrnga and Djeebbana”. *Oceanic Linguistics* 17: 27-37.
- (1979). “Gender and the category unit augmented”. *Oceanic Linguistics* 18 (2): 203-210.
- (1984). “Ndjébbana (Kunubidji) grammar: miscellaneous morphological and syntactic notes”. In: *Papers in Australian Linguistics*. Vol. 16, pp. 119-152. (Pacific Linguistics, A 68). Canberra: Australian National University.
- (1990). “The Addressee: or Is the second person singular”. *Studies in Language* 14 (2): 429-432.
- MEINHOF, CARL (1906). *Grundzüge einer vergleichenden Grammatik der Bantusprachen*. Berlin: Dietrich Reimer.
- MERLAN, FRANCESCA C. (1982). *Mangarayi*. (Lingua Descriptive Studies, 4). Amsterdam: North Holland Publishing Company.
- (1983). *Ngalakan Grammar, Texts and Vocabulary*. (Pacific Linguistics, B 89). Canberra: Australian National University.
- METCALFE, C.D. (1975). *Bardi Verb morphology (Northwestern Australia)*. (Pacific Linguistics, B 30). Canberra: Australian National University.
- MILNE, LESLIE (1921). *An Elementary Palaung Grammar*. Oxford: Clarendon Press.
- MINCH, ANDREW SCOTT (1991). *Essential Elements of Amanab Grammar*. Ph.D. dissertation, University of Texas.
- MITHUN, MARIANNE (1986). “Disagreement: the case of pronominal affixes and nouns”. In: Deborah Tannen & James E. Alatis (eds.) *Language and Linguistics: The Interdependence of Theory, Data, and Applications*, pp. 50-66. Washington: Georgetown University Press.
- (1991). “The Development of Bound Pronominal Paradigms”. In: Winfred P. Lehmann & Helen-Jo Jakusz Hewitt (eds.) *Language Typology 1988: Typological Models in Reconstruction*, pp. 85-104. (Amsterdam Studies in the Theory and History of Linguistic Science, 81). Amsterdam: John Benjamins.
- MIXCO, MAURICIO (1997). *Mandan*. (Languages of the World/Materials, 159). München: Lincom.
- MONOD-BECQUELIN, AUREOLE (1975). *La Pratique Linguistique des Indiens Trumai*. (Langues et Civilisations a Tradition Orale, 9). Paris: Société d'Études Linguistiques et Anthropologiques de France.
- MORAVCSIK, EDITH A. (1978). “Agreement”. In: Joseph H. Greenberg (ed.) *Universals of Human Language*. Vol. 4. Syntax, pp. 331-374. Stanford: Stanford University Press.
- (1988). “Agreement and Markedness”. In: Michael Barlow & Charles A. Ferguson (eds.) *Agreement in Natural Language*, pp. 89-106. Stanford: Center for the study of Language and Information.
- (1994). “Group plural – associative plural or cohort plural”. Unpublished manuscript June 11, 1994. (Message to Linguist List, available at

- <http://linguistlist.org/issues/5/5-681.html>).
- MORPHY, FRANCES (1983). "Djapu, a Yulngu Dialect". In: R.M.W. Dixon & Barry J. Blake (eds.) *Handbook of Australian Languages*. Vol. 3, pp. 1-190. Amsterdam: John Benjamins.
- MOSEL, ULRIKE (1984). *Tolai syntax and its historical development*. (Pacific Linguistics, B 92). Canberra: Australian National University.
- MOUS, MAARTEN (in prep). *Alagwa: grammar, texts and lexicon*.
- MÜHLHÄUSLER, PETER & ROM HARRÉ (1990). *Pronouns and People: the Linguistic construction of social and personal identity*. Oxford: Blackwell.
- MÜLLER, FRIEDRICH (1876-1887). *Grundriss der Sprachwissenschaft*. Wien: Alfred Hölder.
- MUNRO, PAMELA (1976). *Mojave Syntax*. New York: Garland.
- MURANE, ELIZABETH (1974). *Daga Grammar, from morpheme to discourse*. (Summer Institute of Linguistics Publications in Linguistics and Related Fields, 43). Oklahoma: Summer Institute of Linguistics.
- MURATORI, P. CARLO (1938). *Grammatica Lotuxo*. Verona: Missioni Africane.
- MUYSKEN, PIETER (1977). *Syntactic Developments in the Verb Phrase of Ecuadorian Quechua*. Ph.D. dissertation, Universiteit van Amsterdam.
- NABABAN, P. W. J. (1981). *A Grammar of Toba-Batak*. (Pacific Linguistics, D 37). Canberra: Australian National University.
- NEBEL, P.A. (1948). *Dinka Grammar*. Verona: Missioni Africane.
- NETTLE, DANIEL (1999). *Linguistic diversity*. Oxford: Oxford University Press.
- NEWMAN, STANLEY (1967). "Classical Nahuatl". In: Robert Wauchope & Norman A. McQuown (eds.) *Handbook of Middle American Indians*. Vol. 5: Linguistics, pp. 179-200. Austin: University of Texas Press.
- (1980). "Functional Changes in the Salish Pronominal System". *International Journal of American Linguistics* 46 (3): 155-167.
- NICHOLS, JOHANNA (1992). *Linguistic Diversity in Space and Time*. Chicago: University of Chicago Press.
- (1994). "Chechen". In: Riëks Smeets (ed.) *North East Caucasian Languages*, pp. 1-77. (The Indigenous Languages of the Caucasus, 4). Delmar: Caravan Books.
- NICKLAS, THURSTON DALE (1974). *The Elements of Choctaw*. Ph.D. dissertation, University of Michigan.
- NORMAN, JERRY (1988). *Chinese*. (Cambridge Language Surveys). Cambridge: Cambridge University Press.
- NOVELLI, BRUNO (1985). *A Grammar of the Karimojong Language*. (Language and Dialect Studies in East Africa, 7). Berlin: Dietrich Reimer.
- NOYER, ROBERT ROLF (1992). *Features, positions and affixes in autonomous morphological structure*. Ph.D. dissertation, MIT, Boston.
- OATES, WILLIAM & LYNETTE OATES (1964). "Gugu-Yalanji Linguistic and Anthropological Data". In: *Occasional Papers in Aboriginal Studies*. Vol. 2, pp. 1-17. Canberra: Australian Institute of Aboriginal Studies.
- (1969). *Kapau Pedagogical Grammar*. (Pacific Linguistics, C 10). Canberra: Australian National University.
- ODDEN, D. (1983). "Aspects of Didinga Phonology and Morphology". In: M. Lionel Bender (ed.) *Nilo-Saharan Language Studies*. Michigan: African Studies Center.
- OSBORNE, C.R. (1974). *The Tiwi Language*. Canberra: Australian Institute of

Aboriginal Studies.

- PANDHARIPANDE, PAJESHWARE V. (1997). *Marathi*. London: Routledge.
- PARKER, ELIZABETH (1986). "Mundani Pronouns". In: Ursula Wiesemann (ed.) *Pronominal Systems*, pp. 131-166. Tübingen: Gunter Narr.
- PARKER, G.J. (1970). *Southeast Ambrym Dictionary*. (Pacific Linguistics, C 17). Canberra: Australian National University.
- PATON, W.F. (1971). *Ambrym (Lonwolwol) Grammar*. (Pacific Linguistics, B 19). Canberra: Australian National University.
- PAYNE, DAVID L. (1981). *The Phonology and Morphology of Axininca Campa*. (Summer Institute of Linguistics Publications in Linguistics and Related Fields, 66). Arlington: Summer Institute of Linguistics.
- PAYNE, DORIS L. (1982). "Chickasaw Agreement Morphology, a Functional Explanation". In: Paul J. Hopper & Sandra A. Thompson (eds.) *Studies in Transitivity*, pp. 351-378. (Syntax and Semantics, 15). New York: Academic Press.
- PAYNE, THOMAS E. (1993). *The Twin Stories, Participant coding in Yagua Discourse*. (University of California Publications in Linguistics, 120). Berkeley: University of California Press.
- PENCE, ALAN R. (1968). "An Analysis of Kunimaipa Pronouns". *Kivung* 1 (2): 109-115.
- PIKE, KENNETH L. (1954). *Language in Relation to a Unified Theory of the Structure of Human Behaviour*. Glendale: Summer Institute of Linguistics.
- PLANK, FRANS (1985). "Die Ordnung der Personen". *Folia Linguistica* 19: 111-176.
- (1989). "On Humboldt on the Dual". In: Roberta Corrigan, Fred Eckman & Michael Noonan (eds.) *Linguistic Categorization*, pp. 293-333. (Current Issues in Linguistic Theory, 61). Amsterdam: John Benjamins.
- (1996). "Domains of the dual, in Maltese and in general". *Rivista di Linguistica* 8 (1): 123-140.
- PLANK, FRANS & WOLFGANG SCHELLINGER (1997). "The uneven distribution of genders over numbers: Greenberg Nos. 37 and 45". *Linguistic Typology* 1 (1): 53-101.
- (2000). "Dual Laws in (no) Time". *Sprachtypologie und Universalien Forschung* 53 (1): 46-52.
- PLUNGIAN, VLADIMIR (1995). *Dogon*. (Languages of the World/Materials, 64). München: Lincom.
- POLAK, REV. J.E.R. (1894). *A Grammar and a Vocabulary of the Ipuriná Language*. London: Kegan Paul, Trench, Trübner & Co.
- POPJES, JACK & JO POPJES (1986). "Canela-Krahô". In: Desmond C. Debyshire & Geoffrey K. Pullum (eds.) *Handbook of Amazonian Languages*. Vol. 1, pp. 128-199. Berlin: Mouton de Gruyter.
- POPOVICH, HAROLD (1986). "The Nominal Reference System of Maxakali". In: Ursula Wiesemann (ed.) *Pronominal Systems*, pp. 351-358. Tübingen: Gunter Narr.
- POPPE, NICHOLAS (1960). *Buriat Grammar*. Bloomington: Indiana University.
- POSNER, REBECCA (1996). *The Romance Languages*. (Cambridge Language Surveys). Cambridge: Cambridge University Press.
- PRESS, MARGERET LAURITSEN (1975). *A Grammar of Chemehuevi*. Ph.D. dissertation, University of California, Los Angeles.
- PRIESTLY, T.M.S. (1993). "Slovene". In: Comrie Bernard & Greville G. Corbett (eds.)

- The Slavonic Languages*, pp. 388-451. London: Routledge.
- QUIRK, RANDOLPH, SIDNEY GREENBAUM, GEOFFREY LEECH & JAN SVARTVIK (1985). *A Comprehensive Grammar of the English Language*. London: Longman.
- RATH, JOHANN C. (1981). *A Practical Heiltsuk-English Dictionary (with a grammatical introduction)*. Ph.D. dissertation, Universiteit Leiden.
- REED, JUDY & DAVID L. PAYNE (1986). "Asheninca (Campa) Pronominals". In: Ursula Wiesemann (ed.) *Pronominal Systems*, pp. 323-331. Tübingen: Gunter Narr.
- REH, MECHTHILD (1996). *Anywa Language: Description and Internal Reconstruction*. (Nilo-Saharan Linguistic Analyses and Documentation, 11). Köln: Rüdiger Köppe.
- REICHARD, GLADYS A. (1925). *Wiyot Grammar and Texts*. Berkeley: University of California Press.
- (1951). *Navaho Grammar*. (Publications of the American Ethnological Society, 11). New York: J.J. Augustin.
- REID, LAWRENCE A. (ed.) (1971). *Philippine Minor Languages*. (Oceanic Linguistics special publication, 8). Honolulu: University of Hawai Press.
- REINISCH, LEO (1879). *Die Nuba Sprache: Erster Teil: Grammatik und Texte*. Wien: Wilhelm Braumüller.
- (1881). *Die Kunama-Sprache in Nord-Ost Afrika*. Wien: Carl Gerold's Sohn.
- RENCK, G.L. (1975). *A Grammar of Yagaria*. (Pacific Linguistics, B 40). Canberra: Australian National University.
- RICE, KEREN (1989). *A Grammar of Slave*. Berlin: Mouton de Gruyter.
- RICHARD, ERNEST L. (1975). "Sentence Structure of Guhu-Samane". In: Tom E. Dutton (ed.) *Studies in Language of Central and South-East Papua*, pp. 771-816. (Pacific Linguistics, C 29). Canberra: Australian National University.
- RIJKHOFF, JAN & DIK BAKKER (1998). "Language sampling". *Linguistic Typology* 2 (3): 263-314.
- ROBERTS, JOHN R. (1987). *Amele*. London: Croom Helm.
- ROBERTSON, STUART & FREDERIC G. CASSIDY (1954). *The development of modern English*. Englewood Cliffs: Prentice-Hall.
- ROBINETT, FLORENCE M. (1955). "Hidatsa II: affixes". *International Journal of American Linguistics* 21 (2): 160-177.
- RODRIGUES, ARYON D. (1990). "You and I = Neither You nor I: The Personal System of Tupinambá". In: Doris L. Payne (ed.) *Amazonian Linguistics*, pp. 393-406. Austin: University of Texas Press.
- ROJAS CURIEUX, TULIO (1991). "Las estructuras de la oracion en Paez". In: Tulio Rojas Curieux (ed.) *Estudios Gramaticales de la Lengua Paez (Nasa Yuwe)*, pp. 7-100. Colciencias: Universidad de los Andes.
- ROSS, MALCOLM & JOHN NATU PAOL (1978). *A Waskia sketch and vocabulary*. (Pacific Linguistics, B 56). Canberra: Australian National University.
- ROTLAND, FRANZ (1982). *Die Südnilotischen Sprachen: Beschreibung, Vergleichung und Rekonstruktion*. (Kölner Beiträge zur Afrikanistik, 7). Berlin: Dietrich Reimer.
- ROWE, JOHN HOWLAND (1974). "Sixteenth and Seventeenth Century Grammars". In: Dell H. Hymes (ed.) *Studies in the History of Linguistics: traditions and paradigms*, pp. 361-379. Bloomington: Indiana University Press.
- ROYEN, GERLACH (1929). *Die nominalen Klassifikations-Systeme in den Sprachen der Erde: historisch-kritische Studie, mit besonderer Berücksichtigung des*

- Indogermanischen*. Mödling: Anthropos.
- RUDE, NOEL EMERSON (1985). *Studies in Nez Perce Grammar and Discourse*. Ph.D. dissertation, University of Oregon.
- SALMINEN, TAPANI (1998). "Nenets". In: Daniel Abondolo (ed.) *The Uralic Languages*, pp. 516-547. London: Routledge.
- SANTANDREA, STEFANO (1963). *A concise Grammar Outline of the Bongo Language*. Rome: Sodality of St. Peter Claver.
- DE SANTO TOMÁS, FRAY DOMINGO (1560). *Grammatica o arte de la lengua general de los Indios de los Reynos del Peru*. Madrid: Ediciones de cultura Hispanica. (facsimile 1994 with transcription and comments by Rodolfo Cerron-Palomino).
- SANZHEYEV, G.D. (1973). *The Modern Mongolian Language*. (Languages of Asia and Africa). Moscow: Nauka.
- SAPIR, EDWARD (1930). "Southern Paiute, A Shoshonean Language". In: William Bright (ed.) *Southern Paiute and Ute Linguistics and Ethnography*. (The collected works of Edward Sapir, 10). Berlin: Mouton de Gruyter. (Collection published in 1992).
- DE SAUSSURE, F. (1916). *Cours de Linguistique Générale*. Paris: Payot. (Third edition, reprint 1968).
- SAXON, LESLIE ADELE (1986). *The Syntax of Pronouns in Dogrib (Athapascan): some theoretical consequences*. Ph.D. dissertation, University of California, San Diego.
- SCHACHTER, P. & F.T. OTANES (1972). *Tagalog Reference Grammar*. Berkeley: University of California Press.
- SCHAUB, WILLY (1985). *Babungo*. London: Croom Helm.
- SCHEGLOFF, EMANUEL A. (1996). "Some Practices for Referring to Persons in Talk-in-Interaction: A Partial Sketch of a Systematics". In: Barbara A. Fox (ed.) *Studies in Anaphora*, pp. 437-486. (Typological Studies in Language, 33). Amsterdam: John Benjamins.
- SCHMALSTIEG, WILLIAM R. (1998). "The Baltic Languages". In: Anna Giacalone Ramat & Paolo Ramat (eds.) *The Indo-European Languages*, pp. 454-479. London: Routledge.
- SCHMIDT, WILHELM (1926). *Die Sprachfamilien und Sprachkreise der Erde*. Heidelberg: Carl Winter.
- SCHÖNFELD, M. (1959). *Historische Grammatica van het Nederlands*. Zutphen: Thieme. (6th edition).
- SCHULZE-BERNDT, EVA (2000). *Simple and Complex Verbs in Jaminjung. A study of event categorisation in an Australian language*. Ph.D. dissertation, Universiteit Nijmegen.
- SCHUSTER-SEWC, H. (1996). *Grammar of the Upper Sorbian language: Phonology and Morphology*. (LINCOM studies in Slavic Linguistics, 3). München: Lincom.
- SCHWARTZ, LINDA J. (1986). "The Function of Free Pronouns". In: Ursula Wiesemann (ed.) *Pronominal Systems*, pp. 405-436. Tübingen: Gunter Narr.
- SCHWARTZ, LINDA J. & TIMOTHY DUNNIGAN (1986). "Pronouns and pronominal categories in southwestern Ojibwe". In: Ursula Wiesemann (ed.) *Pronominal Systems*, pp. 285-322. Tübingen: Gunter Narr.
- SEILER, WALTER (1985). *Imonda, a Papuan Language*. (Pacific Linguistics, B 93). Canberra: Australian National University.

- SENFT, GUNTER (1986). *Kilivila: The language of the Trobriand Islanders*. (Mouton Grammar Library, 3). Berlin: Mouton de Gruyter.
- SHARPE, MARGARET C. (1972). *Alawa Phonology and Grammar*. (Australian Aboriginal Studies, 37). Canberra: Australian Institute of Aboriginal Studies.
- SHELDEN, HOWARD (1991). "Galela Pronominal Verb Prefixes". In: Tom E. Dutton (ed.) *Papers in Papuan Linguistics*. Vol. 1, pp. 161-175. (Pacific Linguistics, A 73). Canberra: Australian National University.
- SHIBATANI, MASAYOSHI (1990). *The Languages of Japan*. (Cambridge Language Surveys). Cambridge: Cambridge University Press.
- SIEWIERSKA, ANNA (1998). "On nominal and verbal person marking". *Linguistic Typology* 2 (1): 1-56.
- SIEWIERSKA, ANNA & DIK BAKKER (1996). "The Distribution of Subject and Object Agreement and Word Order Type". *Studies in Language* 20 (1): 115-161.
- (2000). "Person asymmetries in the grammaticalization of agreement". Paper presented at Antwerpen: Cognitive Typology (april 14, 2000).
- SILVERSTEIN, MICHAEL (1976). "Hierarchy of features and ergativity". In: R.M.W. Dixon (ed.) *Grammatical Categories in Australian Languages*, pp. 112-171. Canberra: Australian Institute of Aboriginal Studies.
- SIMON, HORST (1999). *Für eine grammatische Kategorie 'Respekt' im Deutschen. Synchronie, Diachronie und Typologie der deutschen Anredepronomina*. Ph.D. dissertation, Humboldt-Universität zu Berlin.
- SIMONS, LINDA (1986). "The Pronouns of To'abaita". In: Ursula Wieseemann (ed.) *Pronominal Systems*, pp. 21-36. Tübingen: Gunter Narr.
- SIMS-WILLIAMS, NICHOLAS (1998). "The Iranian Languages". In: Anna Giacalone Ramat & Paolo Ramat (eds.) *The Indo-European Languages*, pp. 125-153. London: Routledge.
- SMITH, KENNETH D. (1979). *Sedang Grammar: Phonological and Syntactic Structure*. (Pacific Linguistics, B 50). Canberra: Australian National University.
- SMITH-STARK, CEDRIC T. (1974). "The Plurality Split". *Chicago Linguistic Society* 10 (1): 657-672.
- SMOTHERMON, JEFFREY R., JOSEPHINE H. SMOTHERMON & PAUL S. FRANK (1995). *Bosquejo del Macuna: Aspectos de al cultura material de los macunas*. Santafé de Bogotá: Instituto Linguístico de Verano.
- SNEDDON, J.N. (1975). *Tondano Phonology and Grammar*. (Pacific Linguistics, B 38). Canberra: Australian National University.
- SOHN, HO-MIN (1994). *Korean*. London: Routledge.
- SOKOLOVSKAYA, NATALIA K. (1980). "Nekotorye semanticeskie universalii v sisteme licnyx mestoimenij [Some semantic universals in systems of personal pronouns]". In: Ivan F. Vardul' (ed.) *Teorija i tipologija mestoimenij [Theory and typology of pronouns]*, pp. 84-102. Moskva: Nauka. (translations of the universals by E. Filimonova are available at <http://www.ling.uni-konstanz.de:591/Universals/introduction.html>).
- SPAGNOLO, FR. L. M. (1933). *Bari Grammar*. Verona: Missioni Africane.
- SRIDHAR, S.N. (1990). *Kannada*. London: Routledge.
- STAIRS, EMILY F. & BARBARA ERICKSON HOLLENBACH (1969). "Huave Verb Morphology". *International Journal of American Linguistics* 35: 38-53.
- STASSEN, LEON (1997). *Intransitive Predication*. (Oxford Studies in Typology and Linguistic Theory). Oxford: Clarendon Press.
- (to appear). "Some universal characteristics of noun phrase conjunction". In: Frans

- Plank (ed.) *Noun phrase structure in the languages of Europe*. Berlin: Mouton de Gruyter.
- STEINHAUER, HEIN (1985). "Number in Biak: Counterevidence to two alleged language universals". *Bijdragen tot de Taal-, Land- en Volkenkunde* 141: 462-485.
- STEPHENS, JANIG (1993). "Breton". In: Martin J. Ball & Janus Fife (eds.) *The Celtic Languages*, pp. 349-409. London: Routledge.
- STONE, GERALD (1993). "Sorbian". In: Bernard Comrie & Greville G. Corbett (eds.) *The Slavonic Languages*, pp. 593-685. London: Routledge.
- STREITBERG, WILHELM (1920). *Gotisches Elementarbuch*. Heidelberg: Carl Winter.
- STROM, CLAY (1992). *Returã Syntax*. (Studies in the Languages of Colombia, 3). Arlington: Summer Institute of Linguistics.
- SUÁREZ, JORGE A. (1983). *The Mesoamerican Indian Languages*. (Cambridge language Surveys). Cambridge: Cambridge University Press.
- SUÁREZ ROCA, JOSÉ LUIS (1992). *Lingüística misionera española*. Oviedo: Pentalfa Ediciones.
- SUSNIK, BRANISLAVA (1961/62). *Estudio Guayaki, sistema fonético y temático*. Asunción: Museo Etnográfico 'Andres Barbero'.
- (1973). *La Lengua de los Ayoweos-Moros*. Asunción: Museo Etnográfico 'Andres Barbero'. (2nd).
- (1977). *Lengua Maskoy, su hablar, su pensar, su vivencia*. Asunción: Museo Etnográfico 'Andres Barbero'.
- SWADESH, MORRIS (1936). "Nootka Internal Syntax". *International Journal of American Linguistics* 9: 77-102.
- (1946). "Chitimacha". In: Cornelius Osgood (ed.) *Linguistic Structures of Native America*, pp. 312-336. (Publications in Anthropology, 6). New York: Viking Fund.
- SZEMERÉNYI, OSWALD J.L. (1990). *Introduction to Indo-European Linguistics*. Oxford: Clarendon Press.
- TAYLOR, F.W. (1921). *A first grammar of the Adamawa Dialect of the Fulani language*. Oxford: Clarendon Press.
- TEETER, KARL V. (1964). *The Wiyot Language*. Berkeley: University of California Press.
- THELWALL, ROBIN (1983). "Meidob Nubian: Phonology, Grammatical Notes and Basic Vocabulary". In: Lionel M. Bender (ed.) *Nilo-Saharan Language Studies*, pp. 97-113. Michigan: African Studies Center.
- THOMAS, DAVID D. (1955). "Three analyses of the Ilocano pronoun system". *Word* 11 (2): 204-208.
- THOMSON, N.P. (1975). "Magi Phonology & Grammar – Fifty years afterwards". In: Tom E. Dutton (ed.) *Studies in Language of Central and South-East Papua*, pp. 599-666. (Pacific Linguistics, C 29). Canberra: Australian National University.
- THRÁINSSON, HÖSKULDUR (1994). "Icelandic". In: Ekkehard König & Johan van der Auwera (eds.) *The Germanic Languages*, pp. 142-189. London: Routledge.
- TISDALL, ST. CLAIR (1892). *A simplified grammar of the Gujarati language: together with a short reading book & vocabulary*. London: Kegan Paul, Trench, Trübner & Co.
- TODD, EVELYN M. (1975). "The Solomon Language Family". In: S.A. Wurm (ed.) *Papuan Languages and the New Guinea Linguistic Scene*, pp. 805-846.

- (Pacific Linguistics, C 38). Canberra: Australian National University.
- TOMPA, JÓZSEF (1968). *Ungarische Grammatik*. (Janua Linguarum, series practica, 96). The Hague: Mouton.
- TOPPING, DONALD M. (1973). *Chamorro Reference Grammar*. (PALI language texts: Micronesia). Honolulu: University Press of Hawaii.
- TRIER, JOST (1931). *Der Deutsche Wortschatz im Sinnbezirk des Verstandes*. (Germanische Bibliothek, 31). Heidelberg: Carl Winter.
- TRYON, DARRELL T. (1970). *An Introduction to Maranungku (Northern Australia)*. (Pacific Linguistics, B 15). Canberra: Australian National University.
- (1976). “The Daly Family”. In: R.M.W. Dixon (ed.) *Grammatical Categories in Australian Languages*, pp. 673-690. Canberra: Australian Institute of Aboriginal Studies.
- TSERETELI, K.G. (1978). *The Modern Assyrian Language*. (Languages of Asia and Africa). Moscow: Nauka.
- TUCKER, A.N. & M.A. BRYAN (1966). *Linguistic Analyses of the Non-Bantu Languages of North-Eastern Africa*. London: Oxford University Press.
- TUCKER, A.N. & J.TOMPO OLE MPAAYEI (1953). *A Maasai Grammar with Vocabulary*. London: Longmans, Green and Co.
- TUITE, KEVIN (1997). *Svan*. (Languages of the World/Materials, 139). München: Lincom.
- TURNER, PAUL RAYMOND (1966). *Highland Chontal Grammar*. Ph.D. dissertation, University of Chicago.
- TYLER, STEPHEN A. (1969). *Koya: an outline grammar*. Berkeley: University of California Press.
- UHLIG, GUSTAVUS (1883). *Dionysii Thracis Ars Grammatici*. (Grammatici Graeci, 1). Leipzig: Teubner.
- URBAN, GREG (1985). “Ergativity and Accusativity in Shokleng (Gê)”. *International Journal of American Linguistics* 51 (2): 164-187.
- USPENSKY, B.A. (1972). “Subsystems in Language, their interrelations and their correlated universals”. *Linguistics* 88: 53-71.
- VAN VALIN, ROBERT DETRICK, JR. (1977). *Aspects of Lakhota Syntax*. Ph.D. dissertation, University of California, Berkeley.
- VERSTEEGH, KEES (1997). *The Arabic Language*. Edinburgh: Edinburgh University Press.
- VOEGELIN, CHARLES F. (1937a). “Tübatulabal Grammar”. *University of California Publications in American Archeology and Ethnology* 34 (2): 55-190.
- (1937b). “Tübatulabal Texts”. *University of California Publications in American Archeology and Ethnology* 34 (3): 191-246.
- VOORHOEVE, CLEMENS LAMBERTUS (1965). *The Flamingo Bay Dialect of the Asmat Language*. Ph.D. dissertation, Universiteit Leiden.
- (1975). “Central and Western Trans-New Guinea Phylum Languages”. In: S.A. Wurm (ed.) *Papuan Languages and the New Guinea Linguistic Scene*, pp. 345-460. (Pacific Linguistics, C 38). Canberra: Australian National University.
- VOORHOEVE, JAN (1967). “Personal Pronouns in Bamileke”. *Lingua* 17: 421-430.
- VORLAT, EMMA (1975). *The development of English grammatical theory 1568-1737: with special reference to the theory of parts of speech*. Leuven: University Press.
- DE VRIES, LOURENS (1989). *Studies in Wambon and Kombai*. Ph.D. dissertation,

- Universiteit van Amsterdam.
- WALES, KATIE (1996). *Personal pronouns in present-day English*. (Studies in English Language). Cambridge: Cambridge University Press.
- WATERHOUSE, VIOLA (1967). "Huameltultec Chontal". In: Robert Wauchope & Norman A. McQuown (eds.) *Handbook of Middle American Indians*. Vol. 5: Linguistics, pp. 349-367. Austin: University of Texas Press.
- WEBELHUTH, GERT (ed.) (1995). *Government and Binding Theory and the Minimalist Program*. Oxford: Blackwell.
- WEBER, DAVID JOHN (1986). "Huallage Quechua Pronouns". In: Ursula Wieseemann (ed.) *Pronominal Systems*, pp. 333-350. Tübingen: Gunter Narr.
- (1989). *A Grammar of Huallaga (Huánunco) Quechua*. (University of California Publications in Linguistics, 112). Berkeley: University of California Press.
- WEIMER, HARRY & NATALIA WEIMER (1975). "A Short Sketch of Yareba Grammar". In: Tom E. Dutton (ed.) *Studies in Language of Central and South-East Papua*, pp. 667-730. (Pacific Linguistics, C 29). Canberra: Australian National University.
- WELLS, MARGARET A. (1979). *Siroi Grammar*. (Pacific Linguistics, B 41). Canberra: Australian National University.
- WELMERS, W.E. (1976). *A Grammar of Vai*. Berkeley: University of California Press.
- WESTERMANN, D. (1911). *A Short Grammar of the Shilluk Language*. Berlin: Dietrich Reimer.
- (1924). *Die Kpelle Sprache in Liberia*. Hamburg: Dietrich Reimer.
- WESTRUM, PETER N. & URSULA WIESEMANN (1986). "Berik Pronouns". In: Ursula Wieseemann (ed.) *Pronominal Systems*, pp. 37-46. Tübingen: Gunter Narr.
- WHITELEY, W.H. (1958). *A short description of item categories in Iraqw: with material on Gorowa, Alagwa and Burunge*. (East African Linguistic Studies, 3). Kampala: East African Institute of Social Research.
- WHORF, BENJAMIN LEE (1946). "The Milpa Alta Dialect of Aztec". In: Cornelius Osgood (ed.) *Linguistic Structures of Native America*, pp. 367-398. (Publications in Anthropology, 6). New York: Viking Fund.
- WIESEMANN, URSULA (ed.) (1986a). *Pronominal Systems*. Tübingen: Gunter Narr.
- (1986b). "The pronoun systems of some Je and Macro-Je Languages". In: Ursula Wieseemann (ed.) *Pronominal Systems*, pp. 359-380. Tübingen: Gunter Narr.
- WILKINS, DAVID P. (1989). *Mparntwe Arrernte (Aranda): Studies in the Structure and Semantics of Grammar*. Ph.D. dissertation, Australian National University.
- WILLIAMS, MARIANNE (1976). *A Grammar of Tuscarora*. (Garland studies in American Indian linguistics, 18). New York: Garland.
- WILLMS, ALFRED (1972). *Grammatik der südlichen Berberdialekte*. (Afrikanistische Forschungen, 6). München: J.J. Augustin.
- WILSON, DARRYL (1969). "Suena Grammar Highlights". In: *Papers in New Guinea Linguistics*. Vol. 9, pp. 95-110. (Pacific Linguistics, A 18). Canberra: Australian National University.
- WISE, MARY RUTH (1971). *Identification of Participants in Discourse: A study of aspects of form and meaning in Nomatsiguenga*. (Summer Institute of Linguistics Publications in Linguistics and Related Fields, 28). Norman: Summer Institute of Linguistics.
- WOLFF, JOHN U. (1966). *Beginning Cebuano*. New Haven: Yale University Press.
- WOODBURY, ANTHONY (1981). *Study of the Chevak Dialect of Central Yup'ic*. Ph.D. dissertation, University of California, Berkeley.

- WURM, S.A. (1969). "The Linguistic Situation in the Reef and Santa Cruz Islands". In: *Papers in Linguistics of Melanesia*. Vol. 2, pp. 47-105. (Pacific Linguistics, 21). Canberra: Australian National University.
- (1975a). "The East Papuan Phylum in General". In: S.A. Wurm (ed.) *Papuan Languages and the New Guinea Linguistic Scene*, pp. 783-804. (Pacific Linguistics, C 38). Canberra: Australian National University.
- (1975b). "Eastern Central Trans-New Guinea Phylum Languages". In: S.A. Wurm (ed.) *Papuan Languages and the New Guinea Linguistic Scene*, pp. 461-526. (Pacific Linguistics, C 38). Canberra: Australian National University.
- YOUNG, ROBERT & WILLIAM MORGAN (1987). *The Navaho Language*. Albuquerque: University of New Mexico Press.
- DE ZWAAN, JAN DANIEL (1969). *A Preliminary Analysis of Gogo-Yimidjir*. (Australian Aboriginal Studies, 16). Canberra: Australian Institute of Aboriginal Studies.
- ZWICKY, ARNOLD M. (1977). "Hierarchies of Person". *Chicago Linguistic Society* 13 (1): 714-733.

Samenvatting

De paradigmatische structuur van persoonsmarkering

Een wereld vol talen

Op dit moment worden er ongeveer zeventuizend verschillende talen gesproken in de wereld. Je hoeft maar een klein stuk te reizen om al mensen tegen te komen die je niet meer verstaat. Nog duidelijker wordt de talige verscheidenheid bij het beluisteren van de onbegrijpelijke klankenstroom die je tegenkomt in ‘exotische’ talen uit Afrika of Zuidoost-Azië. Echter, bij nadere bestudering blijken de verschillen tussen al deze talen helemaal niet zo groot te zijn. Alle talen op de wereld vertonen op een abstract niveau vele overeenkomsten. Een beschrijving van, laten we zeggen, het Maranungku (een Australische Aboriginaltaal) lijkt veel op een beschrijving van een willekeurige andere taal, zoals het Hixkaryana (een Zuid-Amerikaanse Indianentaal). In beide beschrijvingen komen zaken voor als naamwoorden en werkwoorden, als vraagzinnen, negatie, voorvoegsels, enzovoorts. Natuurlijk is de vorm van de woorden in elke taal weer anders, en natuurlijk verschilt de manier van zinsbouw en woordvorming van taal tot taal, maar de verschillen worden nooit zo groot dat de talen onvergelijkbaar worden. Het spanningsveld tussen de grote verschillen en de opmerkelijke overeenkomsten tussen talen is een centraal thema van taalkundig onderzoek in de afgelopen twee eeuwen. Vooral in de laatste decennia zijn er veel gedegen beschrijvingen van allerlei talen ter beschikking gekomen, die het, nu meer dan ooit, mogelijk maken om een onderzoek te doen naar de talige variatie in de wereld zonder ver te hoeven reizen. In dit proefschrift is een dergelijk onderzoek beschreven. Ik heb gegevens gebruikt van ongeveer vierhonderd verschillende talen, goed verdeeld over de geografische en talige wereld. Alle gebruikte gegevens komen uit gepubliceerde grammaticale beschrijvingen (§ 1.3).

Het onderwerp van dit onderzoek is de verwoording van verwijzing naar de deelnemers in een gesprek. Hoe drukt een taalgebruiker uit dat de actuele spreker wordt bedoeld (bijvoorbeeld in het Nederlands door gebruik van het woord *ik*) of juist de toesprekene (bijvoorbeeld in het Nederlands door gebruik van het woord *jij*)? Wat voor andere vergelijkbare categorieën komen er voor in de talen van de wereld? Het primaire doel van dit onderzoek is een overzicht te geven van de variatie op dit gebied. Deze doelstelling heeft tot gevolg dat de gegevens vaak op een enigszins eclectische manier zijn verzameld: taalfamilies met bijzondere kenmerken heb ik veel intensiever bestudeerd dan andere taalgroepen om zo de grenzen te verkennen van wat talig mogelijk is. Door deze aanpak kan ik nu met grote zekerheid concluderen dat er

geen harde grens is tussen wat wel en wat niet mogelijk is in de structuur van natuurlijke taal. Steeds weer blijkt dat bepaalde taalstructuren vaker voorkomen dan andere, maar dat er is nooit sprake is van een harde breuk tussen ‘gewone’ en ‘uitzonderlijke’ structuren. Er is een continue verbinding tussen de meer en de minder vaak voorkomende structuren. Extrapolerend leidt dit tot de stelling dat in de talige wereld niets onmogelijk is, maar dat bepaalde structuren alleen minder waarschijnlijk zijn.

Persoonsmarkering

De Nederlandse taal biedt allerlei mogelijkheden om als spreker naar jezelf te verwijzen. De meest voor de hand liggende manier is om het woord *ik* te gebruiken. Met de uiting “ik promoveer” geeft de auteur van dit proefschrift aan dat hij promoveert. Een speciale eigenschap van het woord *ik* is dat het noodzakelijkerwijs naar jezelf verwijst. Oftewel, er is geen andere mogelijkheid dan dat het woord *ik* verwijst naar diegene die het woord uitspreekt. Als u, waarde lezer, de zin “ik promoveer” uitspreekt, beweert u niet dat Michael Cysouw promoveert, maar dat u zelf promoveert. In navolging van de linguïst Otto Jespersen worden dergelijke woorden in de taalkunde ‘shifters’ genoemd. Naar wie of wat verwezen wordt met zulke shifters verandert afhankelijk van de situatie waarin ze worden geuit. Andere voorbeelden van shifters zijn woorden zoals *nu* (een tijdsgebonden shifter) of *hier* (een plaatsgebonden shifter). In het voorliggende proefschrift houd ik mij echter alleen bezig met persoonsgebonden shifters, zoals de woorden *ik*, *jij* en *wij*.

In de bovenstaande paragraaf zijn nog een paar persoonsgebonden shifters voorgekomen, namelijk de woorden *auteur* en *lezer*. Wie bedoeld wordt met de woorden *auteur* en *lezer* is de bovenstaande zinnen is afhankelijk van wie de tekst heeft geschreven en wie hem nu leest. In tegenstelling tot de woorden *ik* en *jij* zijn de woorden *auteur* en *lezer* echter niet noodzakelijkerwijs shifters. In de volgende zin, die zou kunnen voorkomen in een literatuurrecensie, zijn de woorden *auteur* en *lezer* geen shifters: “Door gebruik te maken van deze stijlfiguur tracht de auteur de lezer op het verkeerde been te zetten.” In deze zin verwijst het woord *auteur* niet naar de auteur van deze zin zelf, maar naar de auteur van een boek waarover in de recensie wordt gesproken. Deze bedoelde auteur blijft dezelfde persoon als iemand anders deze zin had geschreven. Het woord *auteur* is dus geen shifter in deze context. Zulke woorden, die mogelijk maar niet noodzakelijk een shifter zijn, komen veel voor. Andere voorbeelden zijn het woord *papa*, dat een shifter kan zijn als een vader tegen zijn kind zegt: “Papa gaat even naar boven”, of een eigennaam dat een shifter is als een docent een leerling vraagt: “En wat zou volgens Jan het antwoord zijn?” Er zijn talen in Zuidoost-Azië waar er schijnbaar geen gespecialiseerde persoonsgebonden shifters zijn; alle equivalenten van het Nederlandse *ik* hebben een meer of minder duidelijke zelfstandige betekenis. Dit betekent dat persoonsgebonden shifters niet altijd een gespecialiseerde linguïstische vorm hebben in natuurlijke taal. Natuurlijk kun je in elke taal “ik” zeggen, maar er zijn talen die hiervoor geen gespecialiseerde talige elementen hebben (§ 1.2.1).

In dit proefschrift beperk ik mij tot de gespecialiseerde persoonsgebonden shifters. Zulke talige elementen noem ik ‘persoonsmarkering’ (§ 1.2.2). Aangezien sommige talen geen gespecialiseerde persoonsmarkering hebben, is persoonmarkering volgens de hier gebruikte definitie geen universele grammaticale categorie. Echter, als er persoonsmarkering is, kan die in een taal verschillende vormen aannemen. In het Nederlands is er werkwoordsvervoeging (*loop, loopt, lopen*) en zijn er persoonlijke voornaamwoorden – verschillend in vorm voor onderwerp (*ik, jij, wij*, etcetera), overige zinsfuncties (*mij, jou, ons*, etcetera) en bezitsaanduiding (*mijn, jouw, ons*, etcetera). De werkwoordsvervoeging wordt in het Nederlands slechts rudimentair gebruikt voor persoonsmarkering: er zijn veel minder werkwoordsvormen dan er persoonlijke voornaamwoorden zijn. In andere talen in de wereld speelt de werkwoordsvervoeging een veel belangrijker rol. In de Irokezische talen, die gesproken worden in de Verenigde Staten, zijn er bijvoorbeeld wel zestig verschillende vormen van het vervoegde werkwoord afhankelijk van wie onderwerp en wie lijdend voorwerp is. Om de verschillende vormen van persoonsmarking wereld goed te kunnen vergelijken heb ik in mijn onderzoek zowel gekeken naar persoonlijke voornaamwoorden als naar persoonsvervoeging van het werkwoord (§ 1.2.4).

Paradigmatische structuur

Persoonsmarkeerders zijn geen eenlingen in de structuur van een taal; ze behoren altijd tot een paradigma. Een paradigma is een groep van woorden (of woorddelen) die allemaal dezelfde plaats innemen in de grammaticale structuur van een taal. Het paradigma van de Nederlandse persoonlijke voornaamwoorden (in de functie van onderwerp) wordt gevormd door de woorden *ik, jij, hij, zij, het, wij* en *jullie*. De woorden in een paradigma krijgen hun precieze betekenis in afgrenzing tot elkaar. De betekenis van het woord *ik* wordt mede bepaald doordat er ook een woord *wij* is. Het woord *ik* wordt gebruikt voor “de spreker en geen andere personen” en het woord *wij* heeft de complementaire betekenis “de spreker samen met andere personen”. Deze afgrenzing is niet noodzakelijkerwijs aanwezig in een taal. Bijvoorbeeld in het Salt-Yui, een taal die wordt gesproken in Papua Nieuw-Guinea, is het woord *na* de vertaling van het Nederlandse woord *ik*. Hetzelfde woord *na* wordt echter ook gebruikt in de betekenis van het Nederlandse woord *wij*. De afgrenzing tussen *ik* en *wij* is hier niet aanwezig. In het Mandara, een taal die wordt gesproken in Kameroen, is de situatie precies omgekeerd. In deze taal zijn er meer dan twee woorden voor de betekenissen van de Nederlandse woorden *ik* en *wij*. Het - woord *yá* wordt in het Mandara gebruikt voor “de spreker en geen andere personen” en is geheel equivalent aan het Nederlandse woord *ik*. De woorden *má* en *há* echter zijn allebei vergelijkbaar met het Nederlandse woord *wij*, maar met een speciale wederzijdse afgrenzing. Het woord *má* wordt gebruikt voor “de spreker, de toegesprokene en eventuele andere personen” (omdat de toegesprokene erbij wordt betrokken heet dit ‘inclusief’) en het woord *há* wordt gebruikt voor “de spreker en andere personen, maar niet de toegesprokene” (omdat de toegesprokene wordt uitgesloten heet dit ‘exclusief’). Deze drie talen – Nederlands, Salt-Yui en Mandara – hebben elk hun manier om de begrippen “ik” en “wij” uit te drukken. Ze verschillen niet in wat ze kunnen uitdrukken, maar ze verschillen in hoe ze de per-

soonsmarkering verdelen over verschillende woorden. Deze opdeling noem ik de ‘paradigmatische structuur’ (§1.2.3).

In een taal zijn er meestal een aantal verschillende paradigma’s voor persoonsmarkering. In het Nederlands zijn er drie verschillende paradigma’s van persoonlijke voornaamwoorden en is er één paradigma van de persoonsvervoeging van werkwoorden. Deze paradigma’s hebben elk hun eigen paradigmatische structuur. De structuur van elk van deze Nederlandse paradigma’s heb ik opgenomen in mijn verzameling van paradigmatische structuren, en op dezelfde wijze heb ik gegevens van honderden andere talen in de wereld verzameld.

Enkelvoudige persoonsmarkering

In verreweg de meeste persoonsparadigma’s in de wereld worden minstens drie enkelvoudige categorieën onderscheiden: “spreker” (*ik*, ‘eerste persoon’), “toegesprokene” (*jij*, ‘tweede persoon’) en “anderen” (*hij, zij, het*, ‘derde persoon’). Er is echter een kleine groep paradigma’s die minder onderscheidingen maakt. De Nederlandse werkwoordsvervoeging onderscheidt bijvoorbeeld maar twee vormen in het enkelvoud (*loop* versus *loopt*). Een paradigma met een dergelijke structuur vertoont ‘enkelvoudige homofonie’. Dit verschijnsel is vrij zeldzaam, maar de spaarzame gevallen vertonen veel variatie. Alle theoretisch mogelijke vormen van enkelvoudige homofonie komen voor in de talen van de wereld. Echter, een duidelijke begrenzing van dit verschijnsel is dat enkelvoudige homofonie alleen maar voorkomt in werkwoordsvervoeging. Ik heb geen voorbeelden kunnen vinden van persoonlijke voornaamwoorden met enkelvoudige homofonie (§2.4). De algemene indruk is dat enkelvoudige homofonie een uitzonderlijk grammaticaal fenomeen is. Het komt weinig voor, en als het voorkomt is het nauwelijks aan structurele beperkingen onderhevig. Een dergelijke vrijheid is blijkbaar alleen mogelijk in werkwoordsvervoeging. Persoonsmarkering in de vorm van zelfstandige woorden is voor de taalgebruikers dusdanig betekenisvol dat er geen verwarring tussen de enkelvoudige personen mag ontstaan (§5.4).

Meestal wordt persoon gemarkeerd door het uitspreken van een bepaalde klank, of een combinatie van klanken. Het is echter ook mogelijk een bepaalde persoon te markeren door een zogenaamd ‘nulelement’ te gebruiken, dat wil zeggen door helemaal geen klank te uiten. Deze afwezigheid van markering kan een betekenis krijgen doordat persoonsmarkering in een paradigma georganiseerd is. In wederzijdse afgrenzing tot de wél aanwezige markering krijgt de afwezigheid van markering betekenis. Deze truc kan natuurlijk maar één keer per paradigma worden gebruikt. De Nederlandse werkwoordsvervoeging gebruikt een nulelement voor de eerste persoon (*ik loop-∅* versus *jij/hij/zij/het loop-t*). In paradigma’s mét enkelvoudige homofonie komen zulke nulelementen voor in alle mogelijke personen. In paradigma’s zónder enkelvoudige homofonie daarentegen komen nulelementen bijna uitsluitend voor in de derde persoon. Ook dit is een indicatie van de uitzonderlijke status van enkelvoudige homofonie (§2.6).

Meervoud versus groep

Naast enkelvoud is er natuurlijk ook meervoud. In de meeste grammatica's wordt het woord *ik* geanalyseerd als 'eerste persoon enkelvoud' en het woord *wij* als 'eerste persoon meervoud'. Het begrip 'meervoud' is echter nogal problematisch in de context van persoonsmarkering. Normaal gesproken verwijst een enkelvoud naar één exemplaar van iets, bijvoorbeeld het woord *stoel* verwijst naar één stoel. Een meervoud verwijst naar meerdere exemplaren van dezelfde soort; het woord *stoelen* verwijst naar een verzameling van meerdere stoelen. Echter, deze redenering gaat niet op voor persoonsmarkering. Het woord *ik* verwijst wel naar één exemplaar, namelijk naar één spreker, maar het woord *wij* verwijst normaal gesproken niet naar meerdere exemplaren van dezelfde soort. *Wij* betekent niet "meerdere *ik*-en"; *wij* betekent bijna altijd "ik en een aantal anderen" (behalve in gevallen als *we want more*, unisono geroepen na een goed concert). De groep personen waarnaar het woord *wij* verwijst, bestaat meestal uit meerdere personen (behalve in het geval van pluralis majestatis), maar deze groep is niet simpelweg het meervoud van *ik*. Het meervoud van de derde persoon *zij* is daarentegen wel het echter meervoud van *hij/zij/het*. Het woord *jullie* tenstlotte kan zowel het echte meervoud van *jij* zijn als ook een groep bestaande uit de toegesprokene samen met anderen. Gezien de heterogene betekenis van 'meervoud' bij persoonsmarkering is het beter het begrip 'meervoudige persoonsmarkering' te vervangen door het meer algemene 'groepsmarkering' (§3.3). Er zijn verschillende groepen mogelijk, afhankelijk van de soort personen die de groep vormen. Er blijken vijf verschillende groepen nodig te zijn om de talige diversiteit in de wereld te kunnen beschrijven (§3.4):

- 1+2 alleen spreker en toegesprokene ('minimale inclusief')
- 1+2+3 spreker, toegesprokene en anderen ('uitgebreide inclusief')
- 1+3 spreker en anderen, maar niet de toegesprokene ('exclusief')
- 2+3 toegesprokene(n) en eventuele anderen ('tweede persoon meervoud')
- 3+3 meerdere anderen ('derde persoon meervoud')

De spreker maakt deel uit van de eerste drie van deze groepen. Deze drie groepen samen noem ik daarom het '1^e-persoonscomplex'. Voor de structuur van een paradigma blijkt het van centraal belang te zijn hoe het 1^e-persoonscomplex wordt gemarkeerd. Drie mogelijke structuren van het 1^e-persoonscomplex heb ik eerder al genoemd. Sommige paradigma's hebben helemaal geen gespecialiseerde talige elementen voor deze drie groepen (bijvoorbeeld de bovengenoemde persoonlijke voornaamwoorden van het Salt-Yui), sommige talen hebben één woord voor alle drie groepen samen (bijvoorbeeld het Nederlandse woord *wij*) en sommige talen hebben twee woorden, één voor de minimale en uitgebreide inclusief samen en één voor de exclusief (bijvoorbeeld de persoonlijke voornaamwoorden van het Mandara). Daarnaast is het ook nog mogelijk dat er alleen voor de inclusief een gespecialiseerd element is en dat de exclusief identiek is aan de eerste persoon enkelvoud. Tot slot is het mogelijk dat alle drie de betekenissen van het 1^e-persoonscomplex door drie verschillende talige elementen worden gemarkeerd (§3.5). Naast deze vijf veel voorkomende mogelijkheden zijn er incidentele gevallen van andere structuren (§3.6). De vijf veel voorkomende structuren van het 1^e-persoonscomplex kunnen worden samengevat in twee implicati-

onele stellingen (§3.7). Ten eerste: als er een gespecialiseerde exclusief is, dan is er ook een gespecialiseerde inclusief (exclusief→inclusief). Ten tweede, als er een gedeelde inclusief is, dan is er ook een gespecialiseerde exclusief (gedeeldeinclusief→exclusief).

De markering van de twee overgebleven groepen ('tweede persoon meervoud' en 'derde persoon meervoud') vertoont een samenhang met de markering van het 1^e-persoonscomplex in de vorm van 'verticale homofonie'. Als er identieke vormen worden gebruikt voor de markering van (een deel van) het 1^e-persoonscomplex en één van de andere twee groepen, dan noem ik dat verticale homofonie. Verticale homofonie is als het ware de meervoudige pendant van enkelvoudige homofonie. Een voorbeeld van verticale homofonie is het meervoud van de Nederlandse werkwoordsvervoeging. De werkwoordsvorm van het 1^e-persoonscomplex (*wij lop-en*) is gelijk aan de vorm van de tweede persoon meervoud (*jullie lop-en*) en van de derde persoon meervoud (*zij lop-en*). Opzienbarend is dat verticale homofonie hoofdzakelijk wordt gevonden in werkwoordsvervoegingen die geen onderscheid maken tussen inclusief en exclusief. Dit kan ook worden geformuleerd als een implicatie: als er een inclusief-exclusief onderscheid is in het paradigma, dan is er geen verticale homofonie (§5.3).

Generalisaties

Hoofdstuk 4 geeft een opsomming van alle mij bekende paradigmatische structuren. In totaal heb ik 61 verschillende paradigmatische structuren gevonden in de talen van de wereld, variërend op een continuum dat reikt van zeer veel voorkomende gevallen tot structuren die ik maar in één taal heb gevonden (§4.8). Er blijkt een duidelijke regelmaat te zijn welke structuren meer voorkomen dan andere. Twee hiërarchieën beschrijven de voorkeursstructuren: de Explicietheidshiërarchie en de Horizontale-homofoniehierarchie.

De Explicietheidshiërarchie (Explicitness Hierarchy) is een samenvatting van de samenhang tussen de eerder besproken kenmerken van paradigma's, namelijk de structuur van het 1^e-persoonscomplex, verticale homofonie en enkelvoudige homofonie. Deze kenmerken blijken in een vaste volgorde op te treden in paradigma's. De meest expliciete soort paradigma vertoont geen homofonie en onderscheidt dus alle acht persoonscategorieën (namelijk drie enkelvoudige categorieën en vijf groepen). De eerste onderscheiding die weggelaten kan worden is het verschil tussen de minimale en de uitgebreide inclusief. Het resultaat van deze neutralisatie is een paradigma's met een geünificeerde inclusief en een exclusief (zoals in het Mandara). In een volgende stap op de hiërarchie kan ook het verschil tussen inclusief en exclusief worden weggelaten (zoals in het Nederlands, waar maar één woord *wij* bestaat). Pas als dit gebeurt is, kan verticale homofonie opduiken in het paradigma. Met andere woorden, verticale homofonie komt pas voor in een paradigma als het verschil tussen inclusief en exclusief niet wordt gemarkeerd. Tot slot komt enkelvoudige homofonie pas voor in een paradigma als er ook verticale homofonie is. Dit laatste stadium van de Explicietheidshiërarchie bevat de minst expliciete soort paradigma's. Hier komen paradigma's voor met soms maar twee of drie verschillende vormen, zoals de Nederlandse werk-

woordsvervoeging. Deze hiërarchie kent op alle punten uitzonderingen, maar meer dan 90% van de onderzochte paradigma's heeft een structuur die behoort tot één van de door de hiërarchie beschreven stadia (§5.5).

Op de Explicietheidshiërarchie vindt een soort specialisatie plaats van de markering van deze rol in het gesprek. Er zijn veel manieren om naar een gespreksdeelnemer te verwijzen, bijvoorbeeld door te verwijzen naar uiterlijke kenmerken, sociale status, of door een naam te gebruiken. De persoonsconcepten 'spreker' en 'toegesprokene' hebben echter alleen met de rol in het gesprek te maken, en niet met enige individuele kenmerken van de gespreksdeelnemers. Voor paradigma's in de laagste stadia van de Explicietheidshiërarchie is het begrip 'persoonsmarkering' bijna niet meer van toepassing (waarbij het begrip 'persoon' staat voor de rol van een gespreksdeelnemer in het gesprek). Persoon wordt slechts marginaal uitgedrukt in deze soort paradigma's. Een voorbeeld van een dergelijk marginaal persoonsparadigma is de Nederlandse werkwoordsvervoeging. Er worden nog wel persoonscategorieën onderscheiden (*loop, loopt, lopen*), maar de markering is erg rudimentair. Hoe hoger in de hiërarchie, hoe zuiverder de markering is van de rol in het gesprek, dat wil zeggen 'persoon' (§10.1.5).

Een omslagpunt op de hiërarchie is het al dan niet aanwezig zijn van een onderscheid tussen inclusief en exclusief. Als er geen inclusief-exclusief onderscheid is, is er veel meer mogelijk in een paradigma. Er is over het algemeen meer structurele variatie (§5.7) en het is bijvoorbeeld mogelijk dat er geslachtsmarkering of sociale markering van spreker of toegesprokene optreedt in het persoonsparadigma (§10.2.2). Als er daarentegen wél een inclusief-exclusief onderscheid aanwezig is in het paradigma, dan is de structuur van het paradigma over het algemeen veel gestroomlijnder. Het gaat dan alleen nog maar om het markeren van de rol in het gesprek ('persoon'), niet om het aanduiden van andere kenmerken van een gespreksdeelnemer.

De andere generalisatie is de Horizontale-homofoniehierarchy (Horizontal Homophony Hierarchy). Horizontale homofonie is het verschijnsel dat enkelvoudige en niet-enkelvoudige vormen gelijk zijn. Een voorbeeld is het Engelse *you*, dat zowel de enkelvoudige betekenis "jij" heeft, als de niet-enkelvoudige betekenis "jullie". Het vóórkomen van horizontale homofonie volgt in meer dan 90% van de gevallen een strikte hiërarchie. De eerste stap op deze hiërarchie is dat de derde persoon horizontale homofonie kan vertonen. Als de derde persoon deze homofonie heeft, dan kan ook de tweede persoon horizontale homofonie vertonen. Pas als tweede en derde persoon deze homofonie hebben, dan kan de eerste persoon horizontale homofonie vertonen. De Engelse persoonlijke voornaamwoorden zijn één van de weinige uitzonderingen op deze hiërarchie, omdat de derde persoon geen horizontale homofonie vertoont (*he/she/it* versus *they*), maar de tweede persoon (*you*) wel (§5.6). De Horizontale-homofoniehierarchy volgt de zogenaamde 'persoonshiërarchie': spreker > toegesprokene > ander. Deze persoonshiërarchie komt op meerdere plaatsen in mijn resultaten terug als een belangrijke regulator van paradigmatische structuur (zie ook §5.2, §5.3.3 en §5.5.3).

Zowel de Explicietheidshiërarchie als de Horizontale-homofoniehierarchy hangen samen met de morfologische status van de paradigma's. Paradigma's kunnen bestaan uit zelfstandige woorden (zoals de Nederlandse woorden *ik* en *jij*) of uit gebonden mor-

femen (zoals de Nederlandse werkwoordsvervoegingen *...-t* en *...-en*). Voor beide besproken hiërarchieën geldt dat hoe minder vormen er worden onderscheiden, hoe groter de kans is dat het paradigma bestaat uit gebonden morfemen. En omgekeerd: hoe meer vormen er worden onderscheiden, hoe kleiner de kans is dat het paradigma bestaat uit gebonden morfemen. Deze samenhang hangt waarschijnlijk samen met het feit dat taalgebruikers zich intuïtief bewust zijn van de betekenis van zelfstandige woorden, in tegenstelling tot gebonden morfemen waarvan de meeste sprekers zich niet bewust zijn. Die taalelementen waar gebruikers over kunnen reflecteren, zoals de zelfstandige persoonlijke voornaamwoorden, mogen niet al te gekke vormen van homofonie vertonen (§ 5.8). Een ander aspect in deze context is dat gebonden morfemen zowel voor als achter de werkwoordsstam kunnen voorkomen. Ook hier blijken interessante restricties op de talige variatie te bestaan, die echter in dit proefschrift niet verder konden worden uitgediept (§ 10.2.3).

Speciale aantallen

Een groep bestaat uit meer dan één persoon, maar hoeveel personen er precies in een groep zitten is niet van belang. In sommige talen bestaat er echter ook persoonsmarkering waar het aantal personen wel van belang is. In de beschrijving van deze talen wordt gesproken over speciale elementen voor groepen van twee personen ('dualis'), drie personen ('trialis'), vier personen ('quadralis') of een kleine, maar niet precies gespecificeerde hoeveelheid van personen ('paucalis'). De meest voorkomende categorie is de dualis. De empirische status van de andere categorieën is zwak. Al deze gevallen van speciale aantallen bestaan inderdaad, maar het zijn geen veelvoorkomende structuren. De beschreven gevallen komen allemaal uit Melanesië (grotweg Nieuw-Guinea en omgeving). Bovendien lijkt het erop dat alle gevallen van trialis eigenlijk gevallen van paucalis zijn. Tot slot heeft het enige mij bekende geval van een quadralis duidelijk meer mogelijkheden dan alleen verwijzing naar groepen van vier personen (§ 7.4).

De dualis komt daarentegen relatief vaak voor in persoonsparadigma's en de beschreven gevallen zijn goed verdeeld over de talige wereld. Een problematisch geval is echter een persoonsmarkeerder die wel bekend staat als de 'eerste persoon inclusief dualis'. Een dergelijk element verwijst naar een groep die alleen bestaat uit de spreker en de toegesprokene. Deze groep bestaat uit twee personen en is dus strikt gesproken een dualis. Er is echter een groot aantal gevallen waar deze vorm de enige dualis is in de hele taal. Zulke talen hebben geen dualis voor zelfstandige naamwoorden en ook niet voor andere persoonsmarkeerders. In zulke gevallen heb ik de 'eerste persoon inclusief dualis' niet als een dualis geïnterpreteerd, maar als een speciale soort groep, de zogenaamde 'minimale inclusief'. Dat deze groep uit twee personen bestaat is een bijverschijnsel (§ 3.6.5, § 4.7 en § 7.2). In een aantal andere persoonsparadigma's komt deze 'eerste persoon inclusief dualis' voor in combinatie met andere dualis-persoonsmarkeerders. In zulke gevallen is er sprake van een 'echte' dualis in de persoonsmarkering van de taal in kwestie.

In hoofdstuk 8 wordt een overzicht gegeven van alle ‘echte’ gevallen van een dualis. In totaal heb ik 32 verschillende paradigmatische structuren met een dualis gevonden. Net zoals bij de paradigma’s zonder dualis, is er ook in dit geval een continuum, variërend van zeer vaak voorkomende paradigmatische structuren tot structuren die ik maar in één uitzonderlijk geval heb gevonden (§ 8.9). Er blijkt ook weer een duidelijke regelmaat te zijn in welke structuren meer voorkomen dan andere. Zo is er ook voor de paradigma’s met een dualis een Explicietheidshiërarchie (§ 8.8.2), maar daarentegen geen duidelijke Horizontale-homofoniehierarchie. In het geval van horizontale homofonie is het meer een kwestie van alles of niets: òf geen horizontale homofonie (en dus een complete set dualisvormen), òf overal horizontale homofonie (en dus helemaal geen dualis). Er is een aantal tussengevallen, maar deze gevallen vertonen geen hiërarchie (§ 8.8.3).

Er zijn aanwijzingen dat er nog meer vormen van persoonsmarkering zijn, in de literatuur wel aangeduid als ‘complexe’ of ‘samengestelde’ persoonlijke voornaamwoorden. Dit speciale fenomeen is alleen beschreven voor een regionaal en genetisch beperkte groep van talen in West-Afrika. De precieze functie van deze voornaamwoorden is onduidelijk, maar het schijnt wel om een in hoge mate gegrammaticaliseerd fenomeen te gaan. Nader onderzoek van deze talen is wenselijk (Appendix A).

Historische dynamiek

Taal is onderhevig aan voortdurende verandering. Nieuwe woorden verrijken de taal, oude woorden verdwijnen uit het taalbewustzijn van een jongere generatie. Ook grammaticale structuren veranderen; echter in dit geval gaan de veranderingen vaak veel sluipender en onttrekken zich aan het bewustzijn van de gebruikers. Om een indruk te krijgen van de mogelijke veranderingen in de paradigmatische structuur van persoonsmarkering heb ik een speciale methode gebruikt, die een mix is van synchrone en diachrone taalvergelijking. Het uitgangspunt van deze methode is de aanname dat cross-linguïstische generalisaties, zoals ze in dit proefschrift zijn beschreven, relevant zijn voor taalverandering. Taalverandering heeft een voorkeur voor structuren die binnen de ‘gewone’ variatie vallen. Ik heb daarom de gevonden hiërarchieën (Explicietheidshiërarchie en Horizontale-homofoniehierarchie) geïnterpreteerd als een hypothese voor taalverandering. Om mogelijke taalveranderingen op het spoor te komen heb ik paradigma’s uit nauw verwante talen gezocht die maar minimaal van elkaar verschillen volgens één van beide hiërarchieën. Als het mogelijk blijkt veel van dergelijke voorbeelden te vinden, dan is dat een argument dat de hiërarchieën kunnen dienen als een schets van de mogelijkheden van taalverandering. Als het niet mogelijk blijkt zulke voorbeelden te vinden, wijst dit erop dat de betreffende stap op de hiërarchie niet direct van belang is voor taalverandering (§ 6.2 en § 9.2). De Explicietheidshiërarchie houdt zich goed onder deze hypothese. Voor alle stadia van de hiërarchie heb ik meerdere voorbeelden van verwante paradigma’s gevonden (§ 6.4, § 9.3 en § 9.4). Echter, voor de Horizontale-homofoniehierarchie is dit niet het geval. Alleen het bovenste deel van deze hiërarchie is terug te vinden in taalverandering. Voor de andere stadia van deze hiërarchie lijkt van alles mogelijk: talen springen heen en weer van het ene stadium naar het andere zonder de hiërarchie te volgen (§ 6.3). Het resul-

taat van dit deel van het onderzoek is een web van samenhangende paradigmatische structuren dat een schets geeft van de historische dynamiek van persoonsparadigma's (§9.6).

Conclusie

De belangrijkste conclusie van dit proefschrift is dat het zinnig is om persoonsmarkering te bekijken vanuit een paradigmatisch perspectief. Persoonsmarkering bestaat niet uit losse talige elementen, maar uit samenhangende structuren van elementen. Daarnaast heeft de beslissing om zowel morfologisch zelfstandige als morfologisch gebonden paradigma's te bestuderen een veel completer beeld opgeleverd van de mogelijke paradigmatische structuren. Natuurlijk blijven er nog veel vragen over voor nader onderzoek. Vooral de status van bezittelijke voornaamwoorden en de verschillen tussen de markering van onderwerp en andere zinsfuncties zijn centrale vragen die niet konden worden behandeld. De resultaten van het onderzoek dat ik zojuist heb beschreven zijn echter een belangrijk uitgangspunt om op zoek te gaan naar antwoorden op deze vragen.

Curriculum Vitae

Michael Cysouw is geboren in Nijmegen in 1970. Na de lagere school 'De Klokkenberg' bezocht hij het Stedelijk Gymnasium in Nijmegen waar hij in 1988 zijn eindexamen behaalde met een sterke bèta-gerichtheid. Van 1988 tot 1991 studeerde hij Wiskunde aan de Katholieke Universiteit Nijmegen, maar door de vele andere boeiende vakken die aan de universiteit gegeven worden, taande zijn wiskundige interesse en raakte hij geïnteresseerd in de taalkunde. Van 1991 tot 1994 studeerde hij Algemene Taalwetenschap, afgesloten met een taaltypologische doctoraalscriptie in 1994. Van 1995 tot 2000 was hij verbonden als Assistent in Opleiding aan de afdeling Algemene Taalwetenschap en Dialectologie aan de KUN. Sinds het najaar 2000 werkt hij als wetenschappelijk medewerker aan het Zentrum für Allgemeine Sprachwissenschaft (ZAS) in Berlijn.