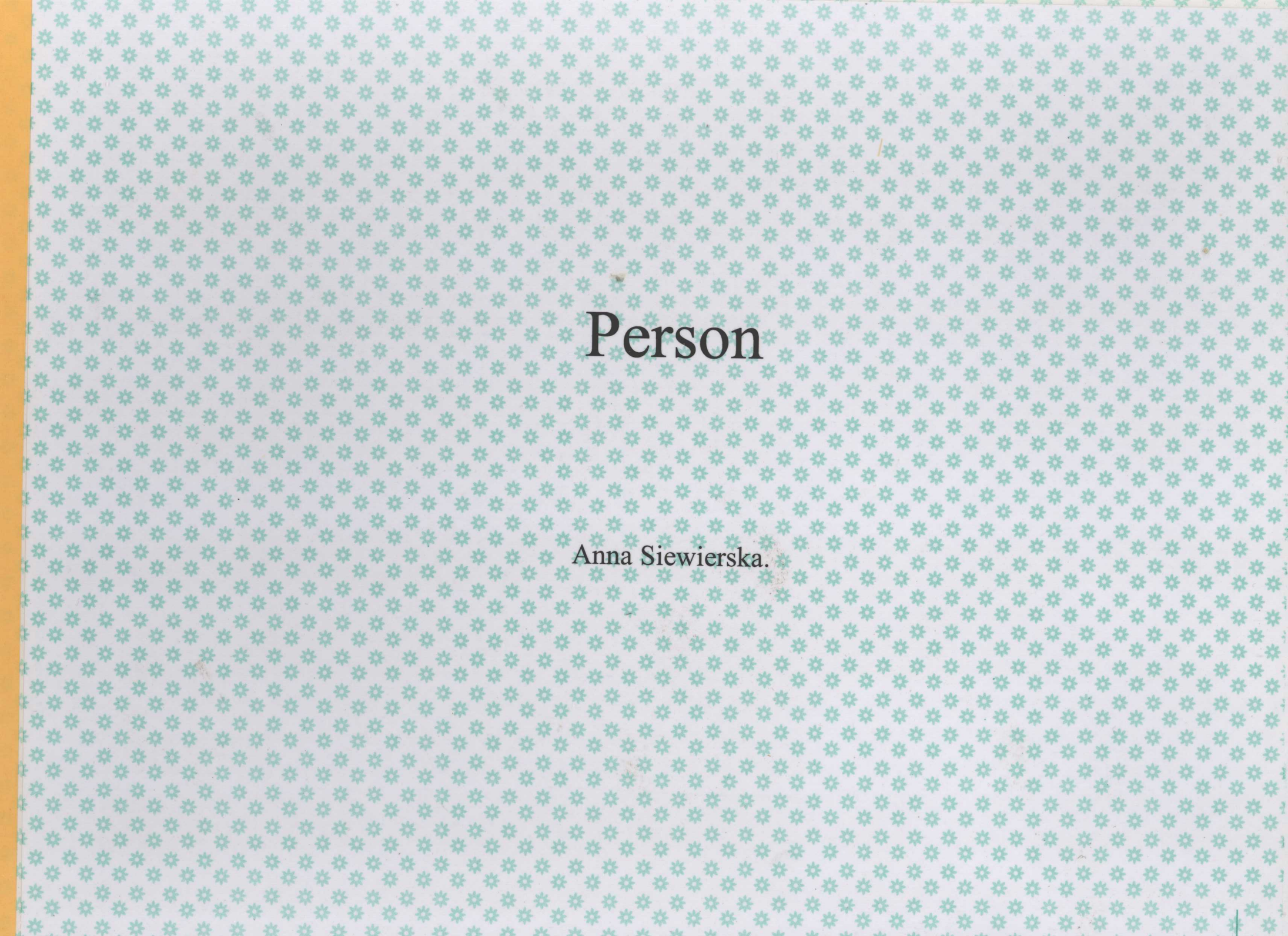


The background of the entire page is a repeating pattern of small, teal-colored, eight-pointed stars or snowflakes arranged in a grid-like fashion on a light cream-colored background.

Person

Anna Siewierska.

CAMBRIDGE TEXTBOOKS IN LINGUISTICS

General editors: S. R. ANDERSON, J. BRESNAN, B. COMRIE, W. DRESSLER,
C. EWEN, R. HUDDLESTON, R. LASS, D. LIGHTFOOT, J. LYONS,
P. H. MATTHEWS, R. POSNER, S. ROMAINÉ, N. V. SMITH, N. VINCENT

Person

江苏工业学院图书馆
藏书章

Contents

<i>List of figures</i>	<i>page</i> xii
<i>List of tables</i>	xiii
<i>Preface</i>	xv
<i>List of abbreviations</i>	xvii
1 Introduction	1
1.1 Person as a grammatical category	1
1.1.1 Person paradigms	4
1.1.2 First and second persons vs third person	5
1.2 The universality of person markers	8
1.3 The nature of this book	14
2 The typology of person forms	16
2.1 Morpho-phonological form	16
2.1.1 Independent forms	17
2.1.2 Dependent person markers	21
2.2 Syntactic function	40
2.2.1 Syntactic function and morpho-phonological form	40
2.2.2 The encoding of syntactic function	47
2.3 Discourse function	67
3 The structure of person paradigms	75
3.1 Fewer than three persons	75
3.2 Variation with respect to number	79
3.2.1 More than one person and the inclusive/exclusive distinction	82
3.2.2 Duals and larger numbers	88
3.2.3 Number and the person hierarchy	92
3.2.4 Towards a typology of paradigmatic structure	96
3.3 Variation in gender	103
3.3.1 Gender and the person hierarchy	104
3.3.2 Gender and number	107
3.3.3 Gender and the inclusive/exclusive distinction	110
3.4 Differences between paradigms	112
3.4.1 Independent vs dependent paradigms	112
3.4.2 Differences between dependent forms	118

4 Person agreement	120
4.1 Anaphoric pronoun vs person agreement marker	121
4.2 The targets of person agreement	127
4.2.1 Predicates	129
4.2.2 Possessed nouns	138
4.2.3 Adpositions and other targets	145
4.3 The controllers of person agreement	148
4.3.1 The person hierarchy	149
4.3.2 The nominal hierarchy	151
4.3.3 The animacy hierarchy	154
4.3.4 The referential hierarchy	156
4.3.5 The focus hierarchy	159
4.4 The markers of person agreement	162
4.4.1 Person agreement and morpho-phonological form	162
4.4.2 The location of person markers	163
5 The function of person forms	173
5.1 Cognitive discourse analysis and referent accessibility	174
5.2 Referent accessibility and the distribution of person forms in discourse	178
5.2.1 Entity saliency	178
5.2.2 Unity	183
5.3 Accessibility and the intra-sentential distribution of person forms	185
5.3.1 Chomsky's Binding Theory	185
5.3.2 Referent accessibility and BT	191
5.3.3 The avoid pronoun constraint	198
5.4 Beyond referent accessibility	200
5.4.1 Long-distance reflexives, logophoricity and point of view	201
5.4.2 Person marker vs other referential expression and speaker empathy	207
5.5 Person markers and impersonalization	210
6 Person forms and social deixis	214
6.1 Alternation in semantic categories	215
6.1.1 Variation in number	216
6.1.2 Variation in person	222
6.1.3 The use of reflexives	224
6.2 Special honorific person markers	228
6.3 Omission of person markers	235
7 Person forms in a diachronic perspective	246
7.1 The sources of person markers	247
7.1.1 Lexical sources	247
7.1.2 Demonstratives	249
7.1.3 Other person markers	251
7.1.4 Conjugated verbal forms	255
7.1.5 Other grammatical markers	260

7.2. From independent person marker to syntactic agreement marker	261
7.2.1 Three accounts of the early stages of the grammaticalization of person markers	263
7.2.2 Syntactic agreement markers	268
7.3 Language externally driven changes in person marking	273
7.3.1 Borrowing of person markers	274
7.3.2 Loss of person agreement	277
Appendix 1. List of languages in the sample by macro-area	282
Appendix 2. Genetic classification of languages cited in the text	284
<i>References</i>	296
<i>Author index</i>	312
<i>Language index</i>	316
<i>Subject index</i>	324

Figures

1 Morphological alignment types in monotransitive clauses	<i>page 51</i>
2 Morphological alignment types in ditransitive clauses	58
3 Singular, vertical and horizontal homophonies in person paradigms	98
4 Relationship between type of agreement markers and type of agreement	126

Tables

2.1 Frequency of lexical and pronominal realization of the s, A and P in six languages	<i>page 41</i>
2.2 Dependent pronominals (as a group) and argument prominence	43
2.3 The alignment of independent and overt dependent person forms	53
3.1 The subdivision of the first-person complex	86
3.2 Major singular and vertical homophonies in Cysouw's sample	99
3.3 Major horizontal homophonies in Cysouw's sample	101
4.1 Distribution of anaphoric and grammatical agreement with different targets	128
4.2 Person agreement in monotransitive clauses relative to alignment	134
4.3 The distribution of person A prefixes vs suffixes relative to basic monotransitive order	165
4.4 The distribution of person P prefixes vs suffixes relative to basic monotransitive order	165
4.5 The order of the A and P relative to each other in prefixal vs suffixal location	167
6.1 Second-person pronoun usage by caste and age in Nepali	230
6.1 Korean person markers	232
7.1 Frequency of first- and third-person pronouns in English according to text type, per million words, based on Biber et al.'s (1999) 40-million-word corpus	267

Preface

This book has two major aims. First of all, it seeks to provide an overview of the various manifestations of the category person in the grammatical system of the world's languages. And secondly it offers a potential account of the principles determining the distribution and form of person markers in utterances. The approach adopted is functional-typological and thus the stress is on the underlying cognitive and discourse basis of person systems and their exponents, on the one hand, and on how these factors are reflected in the existing patterns of cross-linguistic variation, on the other.

While the grammatical category of person is typically associated primarily with that of free personal pronoun, in this book no pride of place is assigned to free as opposed to bound forms or pronouns as compared to agreement markers. A major thread running throughout the discussion is that these different instantiations of the category of person are best viewed as defining both a diachronic and a synchronic cline in regard to their formal and functional properties. Accordingly, no attempt is made to establish universally applicable unique cut-off points on the cline but only to determine the recurring convergences of properties that tend to be found cross-linguistically.

In writing this book I have drawn on the descriptive and analytical insights of numerous scholars. The typological data are taken from over 700 languages. The data originate in the main from reference grammars and grammatical sketches, less frequently from discussions of specific phenomena relating to person marking, both descriptive and theoretical. While I have always made it a point of principle to acknowledge overtly in print each and every source of data or analysis which I have utilized or been inspired by, unfortunately I have not been able to do so in this work. Due to lack of space, I was obliged to eliminate twenty pages of references to the descriptive and theoretical research of my colleagues. Consequently, the list of references at the end of the book contains only the works from which language examples cited in the text are taken and a sub-set of key monographs and articles dealing with various facets of person marking. The full set of publications which I have benefited from is given on my web page <http://www.ling.lang.lancs.ac.uk/staff/anna/person/>.

Over the four years that I have taken to write this book I have had the opportunity to present various aspects of my ideas to colleagues at conferences, workshops and seminars. I am very grateful for all the comments, observations and data that I received. I would like to thank in particular: Mira Ariel, Dik Bakker,

Delia Bentley, Balthasar Bickel, Kirsti Börjars, Dunstan Brown, Bernard Comrie, Grev Corbett, Bill Croft, Martin Haspelmath, Dick Hudson, Peter Kahrel, Marianne Mithun, Johanna Nichols, Frans Plank, Johan van der Auwera, Robert Van Valin, Nigel Vincent and Anne Wichmann. I am especially indebted to the students of the LOT winter-school in Leiden in 2002 who took my course on Person agreement: synchrony and diachrony, for the lively discussion and challenging data which helped me to fine-tune some of my ideas.

I would also like to acknowledge gratefully the support that I received from the Arts and Humanities Research Council (RLS:APN 13302/AN 7261) and Lancaster University as well as from the Max Plank Institute for Evolutionary Anthropology in Leipzig where I did two months' work of data collection, at the invitation of Bernard Comrie and Martin Haspelmath.

In addition I would like to extend my thanks to the team at Cambridge University Press, especially Andrew Winnard, Jacqueline French and Paul Watt.

Finally I would like to thank my friends and family for their support and patience and especially my husband Dik Bakker for agreeing, much more often than he would like, to take second place.

Abbreviations

A	agentive argument of transitive verb
ABL	ablative case
ABS	absolutive case
ACC	accusative
ACT	actual
ADESS	adessive case
ADVR	adverbializer
AFF	affective
AG	agent, agentive
AGR	agreement
AH	addressee honorific
AL	alienable
ALL	allative case
AN	animate
AOR	aurist
APPL	applicative
ART	article
ASP	aspect
ASRT	assertive
ATTR	attributive
AUX	auxiliary
C	common
CAT	catalyst particle
CAUS	causative
CLF	classifier
CLT	clitic
CMP	completive
CNJ	conjunction
COLL	collective
COMP	complementizer
COND	conditional
CONN	connector
CONT	continuous
COP	copula
CORR	correlative

CP	complementizer phrase
CTR	contrast
DAT	dative
DEC	declarative
DEF	definite
DEM	demonstrative
DES	desiderative
DET	determiner
DETR	detransitivizer
DIR	directional
DIST	distal
DR	direct (as opposed to inverse)
DS	different subject
DSTR	distributive
DU	dual
DUR	durative
EMPH	emphatic
EP	epenthetic (vowel or syllable)
ERG	ergative (case)
EVID	evidential
EXCL	exclusive (of the addressee)
F	feminine gender
FACT	factive
FIN	finite
FOC	focus
FUT	future
GEN	genitive case, possessed
GER	gerund
H	honorific
HAB	habitual
HH	high honorific
HORT	hortative
HUM	human
I	inflection
IMP	imperative
IM.PAST	immediate past
IMPF	imperfective
IMPR	impersonal
INAL	inalienable
INAN	inanimate
INCL	inclusive (of the addressee)
IND	indicative
INDEF	indefinite
INDEP	independent

INESS	inessive
INF	infinitive
INST	instrumental
INT.	interrogative
INTR	intransitive
INV	inverse
IO	indirect object
IP	inflection phrase
IPFV	imperfective
IRLS	irrealis
JUS	jussive
LOC	locative
LOG	logophoric
M	masculine gender
MID	middle
MIT	mitigator
MV	main verb
N	noun or nominal
NARR	narrative
NEG	negative element
NF	non-feminine
NFIN	non-finite
NH	non-honorific
NHUM	non-human
NOM	nominative
NP	noun phrase
NPMK	noun phrase marker
NR	nominalizer
NS	nominal suffix
NT	neuter
O	object
OBJ	object
OBL	oblique
OBV	obviative
P	patient argument of transitive verb
PART	participle
PASS	passive
PAST	past
PAT	patient
PAU	paucal
PERF	perfective
PL	plural
PNCT	punctual
POSS	possessive

POT	potential
PP	prepositional phrase
P. PART	past participle
PRED	predicative
PREP	preposition
PRES	present
PRO	pronoun
PROG	progressive
PROX	proximate
PRT	particle
Q	question particle, question word
R	recipient argument of ditransitive verb
REC.PAST	recent past
REDUP	reduplicated form
REFL	reflexive
REL	relative
REM.PAST	remote past
REP	repetitive
RLS	realis
S	single argument of intransitive verb
SEQ	sequential action
SBJ	subject
SG	singular
SPEC	specifier
SS	same subject
STAT	stative
SUBJ	subjunctive
T	theme of ditransitive clause
TAM	tense/aspect/modality marker
TNS	tense
TOP	topic
TRL	trial
TR	transitive
UGR	undergoer
V	verb
VP	verb phrase
1	1st person
2	2nd person
3	3rd person

1 Introduction

The notion of person has been widely discussed in many different fields of study including philosophy, sociology, anthropology, psychology, politics, religion, literature and art. Scholars who have addressed the issue of person within these fields have been concerned with questions such as what is a person, who qualifies as a person, what are the cross-cultural differences in the conceptualization of person, what is the relationship between individual identity and person, how do we identify and reidentify someone other than ourselves, when does a person stop being a person, etc. Though the social and cultural construal of personhood is also a topic of concern within linguistics, particularly sociolinguistics and anthropological linguistics, the notion of person in linguistics is primarily conceived of as a grammatical category, on a par with gender, number, case, tense, etc. Accordingly, it is with person as a category of the grammatical system of languages that this book will be primarily concerned.

1.1 Person as a grammatical category

It is often stated that the grammatical category of person covers the expression of the distinction between the speaker of an utterance, the addressee of that utterance and the party talked about that is neither the speaker nor the addressee. The speaker is said to be the first person, the addressee the second person and the party talked about the third person. This, however, is not quite correct. What is missing from the above characterization is the notion of participant or discourse role. In the case of the first and second persons, the grammatical category of person does not simply express the speaker and addressee respectively, but rather the participant or discourse roles of speaker and addressee.¹ The difference between the two characterizations can be appreciated by comparing the personal pronouns *I* and *you* in (1a) with that of the nominals *mummy* and *Johnny* in (1b).

- (1) a. I will spank you.
b. Mummy will spank Johnny.

¹ This characterization of the grammatical category person draws on the origin of the term person, i.e. mask. Further, it seeks to provide person forms with a sense as opposed to just a reference.

In certain situational contexts, speakers may refer to themselves and their addressees by their proper names, the relations of kinship that they bear to each other, their titles or occupational roles, etc. Thus in (1b) the word *mummy* could be used by a mother with reference to herself and the name *Johnny* with reference to the child whom she is addressing. In such a case, the words *mummy* and *Johnny* can be said to express the speaker and addressee but they cannot be said to express the discourse roles of speaker and addressee as there is nothing in the words *mummy* and *Johnny* to suggest that they are the speaker and addressee respectively. Conversely, this is precisely what is achieved by the two pronouns *I* and *you* in (1a). *I* is always used to refer to the speaker and *you* to the addressee.² Unlike *mummy* and *Johnny*, the two pronominals cannot have any other referents. Moreover, they do not express anything other than that their referents bear the discourse roles of speaker and addressee respectively. Accordingly, only *I* and *you* and not *mummy* and *Johnny* are expressions of the first and second persons. *Mummy* and *Johnny* are lexical expressions which may be used to refer to the speaker and addressee respectively.

In principle, there is no limit to the nature of the lexical expressions that a speaker may use to refer to herself. By contrast, it would be dysfunctional for languages to have a wide range of expressions to denote the discourse roles of speaker, addressee and third party. And indeed they tend not to. The vast majority of the languages of the world have a closed set of expressions for the identification of the three discourse roles embracing the category of person. The special expressions in question are typically called personal pronouns, or even just pronouns. (The word pronoun without additional qualification is generally interpreted as denoting pronouns expressing person.) In this book, however, we will use the terms *person marker* and *person form* in preference to *pronoun*, as the term pronoun is open to a number of interpretations and even under the most liberal of these, not all grammatical markers of the category person are uncontroversially pronominal. More about the notion of pronoun will be said in section 1.2.

Although the grammatical category of person involves only the three-way distinction of speaker, hearer and third party, this does not mean that languages typically have only three person markers. English, which clearly has many more than three person markers, is by no means exceptional. In fact, despite the array of person markers that English has, it does not qualify as a language rich in person markers. Other languages have many more. For instance, Fijian is said to have as many as 135 person forms. There are also languages with considerably fewer person markers than English. Madurese, an Austronesian language now mainly spoken in Java, has only two, *sengkoq* 'I/me' and *tang* 'my'. For the second and third persons, words meaning 'metaphysical body/spirit' and 'sole/alone' accompanied by a definite marker are used.

² This is not quite correct. The second-person form *you* in English, and also in many other languages, has an impersonal or generic use, illustrated in (13b) further below and discussed in more detail in chapter six.

The differences in person-marker inventories found cross-linguistically are in part a reflection of the nature of the grammatical categories in addition to person that the person markers encode. Person markers rarely mark person alone. The grammatical category most closely connected with person is that of number. Two other grammatical distinctions regularly expressed together with person are gender and case. Thus, for example, the English *she* encodes third person, singular number, feminine gender and nominative case, that is the case of the subject. Further grammatical categories which may also be marked together with person include definiteness, obviation, tense, aspect, mood and polarity. The last of these is to be found in the person markers of the Australian language Worora, for example, which, as shown in (2), has a distinctive set of forms used in negative utterances.

(2) Worora (Love 2000:17)

		positive	negative
1SG		ɲaiu	'ɲaiu
2SG		ɲundju	'ɲungi
3SG ³	M	'indja	'kaui
	F	'nijina	'njungi
	NT	'wuna	'kui
	NT	'mana	'maui

In addition to other grammatical categories, person markers may also encode information pertaining to their referents, for example, the social status of the referent vis-à-vis the speaker, their location relative to the speaker or addressee or, much more rarely, their kin relationship and/or generation level. A celebrated instance of person forms reflecting generation levels comes from Lardil (Hale 1966), another Australian language. In Lardil, in the dual and plural, one set of person forms is used for persons who belong to the same generation level or are two levels apart, and a different set of forms for persons one or three generations apart. Thus the form of the second-person dual 'you two' when used to refer to, say, a brother and sister or a grandparent and their grandchild is *kirri*, but when used to refer to a parent and child or great-grandparent and their grandchild is *nyiinki*. More complex systems involving not only considerations of generation level but also of membership within a given moiety (i.e. a particular set of kin categories) are found in other Australian languages, such as Arabana-Wangkangurru (Hercus 1994:117), Adnyamathanha and Kuyani (Schebeck 1973). In these last two languages there are twelve different sets of person markers to mark the kinship associations of the people to whom the person forms refer and, in some instances, also the speaker's relationship to these people. In contrast to the Australian languages mentioned above, in the Tibeto-Burman language Dhimel (King 2001) there are special person forms just for the first- and second-person singular which are reciprocally used only between two distinct groups, one being the parents of

³ Membership in the two sets of neuter forms in Worora, the *wuna* set and the *mana* set is lexically determined.

a husband and a wife and the other, a man and his wife's senior relatives. In exchanges between these two groups the first-person singular is *kya* which contrasts with the typical *ka*, and the second-person singular is *nya* rather than *na*. Yet another factor, in part relating to referents, which has been noted to be encoded in the person markers of a language is speech style. Jacquesson (2001:123) reports that in several dialects of Tiddim, a Tibeto-Burman language, there are two sets of verbal person markers for all three persons: a prefixal set and a suffixal set. The former is used in narratives, the latter in everyday speech.

The other major source of differences in person-marker inventories is variation in morpho-phonological form. In some languages all the person markers are independent words, while others, in addition to such forms, also have person clitics and/or affixes and/or covert, that is zero forms. Bulgarian, for example, apart from independent forms, which may be used for all syntactic functions, has clitics used for objects, and affixes (fused with tense/aspect) used for subjects. All three forms occur in (3).

- (3) Bulgarian (Dimitrova-Vulchanova & Hellan 1999:490)
 Na Ivana kniga-ta az mu= ja= dadox
 to Ivan book-DEF I 3SG:DAT 3SG:ACC give:1SG
 'I gave the book to Ivan.'

1.1.1 Person paradigms

The person markers found in languages do not occur in isolation but rather in closed sets called paradigms. Simplifying somewhat, a paradigm is a set of linguistic expressions that occur in the same syntactic slot in the language. Moreover, each member of a paradigm is in complementary distribution with every other member of the same paradigm. Thus the English person forms *I/you/he/she/it/we/you/they* constitute one paradigm, as each may occur as the subject of an utterance and the use of any one form excludes the possibility of using any of the others (apart from coordinations). The person forms *me/you/him/her/it/us/you/them* belong to another paradigm, since they are employed as objects and complements of prepositions but, crucially, not as subjects. And the forms *my/your/his/her/its/our/your/their* make up a third paradigm used as attributive possessors. In addition to performing the same syntactic function, the members of a single paradigm are also assumed to have the same morpho-phonological form, that is to be all independent forms, or all clitics or affixes, etc.⁴ Consequently, independent and clitic forms, such as the Bulgarian third-person masculine object forms *jemu* and *mu*, are seen as belonging to two different paradigms.

While there are languages which have only one paradigm of person markers used for all syntactic and discourse functions, most languages have several. An

⁴ Occasionally differences in morpho-phonological form are found within what is considered to be a paradigm. For instance, in Fur (Jakobi 1990:28), the dependent object person markers in the singular are suffixes, in the plural clitics or weak forms.

important point to remember is that in those languages which have more than one paradigm of person markers, the structure of the different paradigms need not be the same. The paradigms may differ in regard to the person, number, gender and other distinctions marked. A particularly clear example of such differences between paradigms is presented in (4) from Vinmavis, an Oceanic language spoken on the island of Malakula, in which the independent person markers evince a singular/plural distinction and an inclusive/exclusive one (see section 3.2.1), while the subject prefixes exhibit an opposition between the singular, dual and plural but no inclusive/exclusive contrast.

- (4) Vinmavis (Crowley 2002b:640, 644)
- | Indep form | | Subject prefix (non-future) | |
|------------|-------|-----------------------------|----------|
| 1SG | no | 1SG | nV- /na- |
| 2SG | gu | 2SG | u- |
| 3SG | i | 3SG | i- |
| 1PL INCL | get | 1DU | er- |
| 1PL EXCL | gemem | 2DU | ar- |
| 2PL | gem | 3DU | ar- |
| 3PL | ar | 1PL | it- |
| | | 2PL | at- |
| | | 3PL | at- |

Although the number of distinct person forms in the two paradigms is actually the same, seven (due to the homophony between the second and third persons in both the dual and plural, in the case of the subject prefixes), they differ radically in their internal structure. The existence of such differences makes it difficult to discuss the person system of a language as a whole. Linguists are often tempted to make general statements about the nature of person marking in a language. Such general statements, however, are possible only for some languages, but definitely not others. This has to be kept in mind while reading this book. Just because a particular language is cited as displaying a particular property or feature in some person paradigm, this does not mean that the same holds for all the person paradigms.

1.1.2 First and second persons vs third person

It is generally acknowledged that "there is a fundamental, and ineradicable, difference between the first and second person, on the one hand, and the third person on the other" (Lyons 1977:638). One manifestation of this difference is that whereas the first and second persons are regularly referred to essentially only by person markers, reference to the third person can be achieved by any lexical expression. It should therefore be unsurprising that languages may have first- and second-person markers but no third-person ones. In many of the languages which lack person markers for the third person, demonstrative pronouns corresponding to the English *this* and *that* are used in lieu of third-person markers. This is the case, for instance, in Basque, Comanche, Imbabura Quechua, Lak,

Lavukaleve and Maricopa. In other languages reference to the third person is achieved only via full nominal expressions. This appears to be so, for instance, in Salt Yui (Irwin 1974:32), a Papuan language, where third person is indicated by a specific noun, such as *yai* 'male' or *al* 'female' followed by a demonstrative. There is no distinction in number. Thus *yai i* denotes 'he' or 'they masculine' and *al i*, 'she' or 'they feminine'. And in yet other languages either full nominal expressions are used or, alternatively, no overt expression at all, the absence of an overt expression being interpreted as denoting third person.

A difference between first and second persons as opposed to the third may also be manifested in languages which have person markers for all three persons. Often the forms of the first and second persons are quite different from that of the third. As (5) illustrates, this may be observed in Nosu, a Tibeto-Burman language belonging to the Northern Yi group, spoken by over two million people in Sichuan and northern Yunnan, China.

- (5) Nosu – Northern Yi (Bradley 1993:185)
- | | |
|-----|--------------------|
| 1SG | ŋa ³³ |
| 2SG | nwa ³³ |
| 3SG | tsʰz ³³ |

There may also be a difference in the order of third-person forms as compared to that of first- and second-person forms. For example, in Takale and Gamale, two dialects of the Tibeto-Burman language Kham (Watters 1993:105), when the agent is first or second person, the agent forms precede the patient forms. But when the agent is third person, the agent forms follow the patient ones. Compare (6a,b) with (6c).

- (6) Gamale (Watters 1993:107)
- | | |
|----|-------------------------|
| a. | Nə-hnə-kəŋ-khě |
| | 2SG(A)-look-1SG(P)-PAST |
| | 'You looked at me.' |
| b. | Ye-hnə-rə |
| | 1SG(A)-look-3PL(P) |
| | 'I looked at them.' |
| c. | Ya-hnə-kəŋ-wo |
| | PAST-look-1SG(P)-3SG(A) |
| | 'He looked at me.' |

Another, not uncommon, difference between first and second persons as opposed to third person involves case marking. Third-person forms may take a different set of case markers than first- and second-person forms. For instance, in the Australian language Wambaya (Nordlinger 1998), there are three separate third-person forms, one for the *s* (sole argument of an intransitive clause), another for the *A* (agentive argument of a transitive clause) and a third for the *P* (patient-like argument of a transitive clause). But there are only two forms for the first and

second persons, one for the *s* and *A*, and another for the *P*. Particularly frequent are differences between the first and second persons as compared to the third in regard to number and gender. Number distinctions are often neutralized in the third person, while gender is rarely manifested by second- and hardly ever by first-person forms. (This is discussed in detail in chapter three.)

All of the above differences are typically seen to be a consequence of the fact that first- and second-person forms are inherently deictic expressions, that is their interpretation is dependent on the properties of the extralinguistic context of the utterance in which they occur. Although the first person is always the speaker of the utterance and the second the hearer, the actual identity of each depends on who utters the utterance that contains them to whom, when and where. They belong to the class of expressions often referred to as shifters (Jakobson 1971). Third-person forms, on the other hand, are essentially anaphoric expressions. Their interpretation depends not on the extralinguistic but on the linguistic context of the utterance.⁵ The referent of *he* or *she* is typically established by the preceding discourse, as in (7) or, less often, by the following discourse as in (8).

- (7) – There's no sign of John.
 – He must have missed his train again.
- (8) – She is late again.
 – You mean Sally.
 – Yes.

Third-person forms may be used deictically, as when someone says (pointing to a grinning child who has just been given an enormous ice cream) *He's happy*. Their anaphoric use is, however, the basic one. In fact in some languages, third-person forms can only be used anaphorically, deictic reference being achieved via demonstratives. Much less frequently, in addition to demonstratives there are two sets of third-person forms, one set for deictic reference, and another for anaphoric. This is so in Udihe (Nikolaeva & Tolskaya 2001:753–4), a Tungusic language spoken by about a hundred people in the Russian Far East.

In the linguistic literature, mention is sometimes made not only of the first, second and third persons but of a fourth person. This label is applied to several quite different kinds of categories. For instance, in the French grammatical tradition the term fourth person is often used for the first person plural. In Amerindian studies, especially of Algonkian languages, the label fourth person is used with reference to a less important third person, called an obviative as opposed to a proximate. And in discussions of anaphoric relations across clauses, the term fourth person is used for special third-person forms that indicate coreference, which are also termed logophoric or long-distance reflexives. Under none of the above uses does the fourth person qualify as a bonafide additional discourse category. Therefore, I see no reason for using the term here.

⁵ In place of the terms deictic and anaphoric, some linguists use the terms exophoric and endophoric.

Whereas some scholars seek to expand the number of categories comprising person from three to four, others seek to reduce it. As mentioned earlier, of the three persons only the first and second persons are actual participants in the speech act realized by the utterance containing them. The third person is a not a participant of the speech act. Some linguists, most notably Benveniste (1971), argue that the grammatical category of person should therefore be seen as embracing only the first and second persons with the third person being a non-person. This is not the view adopted in this book. While fully acknowledging the distinctive nature of the third person relative to the first and second, I see no advantage in excluding the third person from the category of person, particularly in a cross-linguistic study such as this one. In fact, as will become apparent in the course of our discussion, doing so would severely skew our understanding of a number of facets of the category of person.

1.2 The universality of person markers

Despite statements such as the following by Benveniste (1971:225) "A language without the expression of person cannot be imagined", the universality of person as a grammatical category is sometimes called into question. The issue of whether all languages display the grammatical category of person is inherently tied to the issue of whether all languages have the category of personal pronoun. What constitutes a personal pronoun is in turn a matter of considerable controversy. The notion of pronominality has been and continues to be discussed in several different contexts and thus the features taken to be characteristic of pronouns are very much dependent on what they are being compared with or opposed to. Traditionally, personal pronouns have been opposed to nouns or NPs. Within the generative approach, ever since Chomsky's (1981) binding conditions, they have been contrasted primarily with anaphors (reflexives). Another line of inquiry opposes personal pronouns to person agreement markers (e.g. Bresnan & Mchombo 1987). And yet other studies seek to characterize pronouns in terms of a scale of structural deficiency (Cardinaletti & Starke 1999). We will have cause to consider all of the above at various points in our discussion, but for the time being, let us just concentrate on the pronoun vs noun distinction.

Traditionally, a personal pronoun is taken to be a morpho-syntactic category, which may be used to substitute for nouns or rather NPs, but differing from the latter in its morphological and syntactic properties. Under this traditional approach various languages, most notably South-east Asian languages such as Thai, Burmese, Vietnamese and Japanese, have been argued to lack personal pronouns, since the expressions used to indicate person display properties of nouns.⁶

⁶ Other languages, such as the Salishan Northern Straits Salish (Jelinek 1998) and Halkomelem (Wiltschko 2002) have been argued to possess only bound pronouns. Such languages will be discussed in section 2.1.1.

More recently, however, what constitutes a pronoun has come to be viewed somewhat differently. In the generative literature (e.g. Noguchi 1997, Bresnan 2001b) a pronoun is seen to be not a morpho-syntactic category but rather a feature that sets off certain lexical items from others. The relevant feature is referential dependency; although pronouns are used to refer to individuals and entities, the identity of their referents can be determined only by the extralinguistic context (for first- and second-person forms) or typically the linguistic context (for third-person forms) or inferentially. This referential deficiency distinguishes them from both proper nouns, which are capable of identifying a referent by themselves, and common nouns, which are semantic predicates requiring a determiner to enable them to be used as referential expressions. In terms of this approach, all or some of the South-East Asian languages mentioned above are seen to have pronouns, but differing in syntactic category from the pronouns in, say, English. English pronouns are treated as determiners, and Japanese pronouns as nouns. The morpho-syntactic differences between the relevant forms in the two languages are thus seen to follow from differences in their categorial status but not in their pronominal status.

In the functional literature, in turn, pronouns in the main continue to be viewed as a morpho-syntactic category but often the distinction between pronoun and noun is considered to be not discrete but scalar, with some pronouns exhibiting less prototypically pronominal and more nominal characteristics than others. This position is most clearly articulated by Sugamoto (1989), who posits the characteristics in (9) as representing the pronominal extreme of what she calls the pronominality scale:

- (9) a. closed class membership
- b. lack of morphological constancy
- c. lack of specific semantic content
- d. lack of stylistic and sociolinguistic implicative properties
- e. expression of grammatical person
- f. inability to take modifiers
- g. restrictions on reference interpretation

These criteria can be used to place person markers on a pronominality scale both across languages and also within languages. For example, if applied to the personal pronouns in English, Polish, Japanese and Thai, the Polish personal pronouns emerge as more pronominal than the English, both as considerably more pronominal than the Japanese forms, and the Japanese forms as more pronominal than those in Thai, as exemplified on the pronominality scale in (10).

- (10) The pronominality scale
- | | |
|--------------------|---------------------|
| + Nominal | + Pronominal |
| ←-----→ | |
| Thai Japanese | English Polish |

Let us first consider the English personal pronouns.

English personal pronouns have most of the properties in (10). They belong to a closed class and, unlike nouns, are not morphologically transparent as far as number or case is concerned. Whereas number with most nouns is indicated by suffixation of /-(ə)z/ (e.g. *dog* vs *dogs* or *dress* vs *dresses*), with pronouns it is indicated by suppletion of the stem (e.g. *I* vs *we*). And whereas nouns may be marked for the genitive case by /'s/ (e.g. *mother's friend*), pronouns again have separate forms (e.g. *I* vs *my*). Further, pronouns convey no semantic content other than that of the grammatical features which are associated with them and do not vary stylistically, while nouns may do so (e.g. *mother* vs *mummy* vs *mum*). And clearly pronouns distinguish between the first, second and third persons, while nouns are necessarily third person. English personal pronouns can, however, co-occur with some of the modifiers that are found with nouns. The plural forms may be modified by a low numeral (e.g. *us two*, *we four*), the accusative forms may be modified by certain adjectives, such as *poor*, *kind*, *evil*, *lucky*, (e.g. *poor me*, *lucky you*) and the nominative forms may be modified by a non-restrictive relative clause, as in *I, who have nothing*, *he who strives*, *wins*. As for reference, the personal pronouns are clearly restricted in regard to their referential interpretations in the sense outlined above. While nouns may be used for both definite (e.g. *the book*, *this book*, *my book*) and indefinite reference (e.g. *a book*, *some book*), personal pronouns are (with few exceptions) definite.⁷ This is reflected in the fact that they do not normally occur with any of the determiners, i.e. articles (e.g. **the he*), demonstratives (e.g. **this she*) or genitives (**my he*) which transform a noun into a definite referential expression and are normally incompatible with the indefinite article (e.g. **a she*). The qualification normally is necessary in view of examples such as those in (11), taken from Noguchi (1997:778–9).

- (11) a. This is not the real me.
 b. Do you know the real you?
 c. That's not a he; that's a she.
 d. It's a he!

The use of the definite article with personal pronouns as in (11a,b) is highly restricted; for most speakers the personal pronoun must be the accusative singular form and, for some, even just the first and second persons (*?This is not the real him.*), and the adjective must be present (**This is not the me.*). The indefinite article is possible only with the third-person nominative (**It's a him*). Such usage thus cannot be seen as actually undermining the essentially definite nature of the personal pronouns.⁸ The above notwithstanding, English personal pronouns are not always used strictly referentially, that is to refer to concrete entities or individuals. For example, in (12) *Kate* has no specific person in mind and thus *he* is used non-referentially.

⁷ For a discussion of reference and definiteness see, e.g., Lyons (1977:177) or Allan (2001:59, 440). The issue will be resumed in sections 4.1 and 4.3.4.

⁸ One way in which such atypical co-occurrences of the article and personal pronoun are dealt with is by assuming that a category conversion has taken place, from a pronoun to a noun.

- (12) Kate still wants to marry a Swede. The problem is that he has to be rich and there are not many rich Swedes around.

In (13) the third-person plural form *they* and the second-person form *you* are used for generic or arbitrary reference (i.e. impersonally).

- (13) a. They say that time heals all pain.
 b. You add the eggs to the butter not the other way round.

And in (14) both the personal and possessive forms are bound by the operator *every* and are thus construed as bound variables.

- (14) a. Every man thinks that he is clever.
 b. Everyone loves his mother.

The personal pronouns in Polish, are very much like their English counterparts. They clearly do form a paradigm, are not transparent morphologically, exhibit restricted possibilities in regard to modification, among which modification by a demonstrative is not included. However, they are necessarily definite. Unlike the English forms, they cannot be used non-specifically, generically or construed as bound variables. For example, an overt personal pronoun as in (15a) can be interpreted as coreferential only with some entity outside of the clause, not as bound by the quantified subject NP of the main clause.

- (15) Polish
 a. Każda kobieta uważa, że ona jest mądra
 every woman considers that she be:3SG:PRES clever
 'Every woman_i thinks that she_i is clever.'
 b. Każda kobieta uważa, że jest mądra
 every woman considers that be:3SG:PRES clever
 'Every woman_i thinks that she_i is clever.'

A bound variable reading is possible but of the person inflection on the verb, that is in the absence of an overt pronoun, as in (15b).⁹ The same holds for a non-specific reading of the Polish equivalent of (12) and a generic interpretation of the Polish version of (13). Both are possible but only in the absence of an overt personal pronoun. Thus if necessary referentiality is viewed as an indicator of greater pronominality, then the Polish personal pronouns are more pronominal than the English ones.¹⁰ The only nominal feature that they do display is the

⁹ In the generative literature the difference between languages like English and Polish in regard to the bound variable interpretation of overt pronouns is attributed to the Overt Pronoun Constraint which is: overt pronouns cannot receive a bound variable interpretation in situations where a null pronoun could occur.

¹⁰ This need not be the case. For instance, Noguchi (1997) does not view the possibility of being used non-referentially as pertinent to pronominality. For him whether a pronoun can be construed as a bound variable is definitive of its determiner as opposed to noun status. Thus under this analysis, English pronouns are more determiner-like, Polish pronouns more noun-like, counter to the ordering on the pronominality scale in (10).

presence of sociolinguistic implications. Among adults, the third-person forms are considered to be informal (see discussion in ch. 6, section 6.1.2).

Turning to the Japanese personal pronouns, there are quite a few variants used to express each of the three persons, carrying different stylistic and sociocultural implications. Some of the existing forms are presented in (16).

- (16) Japanese
 1SG watasi, watakushi, ore, temae, boku, etc.
 2SG anata, kimi, omae, temae, etc.

Unlike in English and Polish, the pronouns do not differ morphologically from nouns, that is they form the plural by the same means as nouns (*tomodachi-tachi* 'friend-PL', *watashi-tachi* 'I-PL') and take the same postpositional case markers. They also display a greater range of modification possibilities and fewer restrictions on the modifiers that they permit than in the case of English or Polish. They may be modified by any adjective (17a), and significantly be preceded by a possessive pronoun (17b) or a demonstrative pronoun (17c).

- (17) Japanese (Noguchi 1997:777)
 a. tiisai/sinsetuna/ookii kare
 small /kind/big he
 '*small/kind/big he'
 b. watsi-no kare
 I-GEN he
 '*my he.' (=boyfriend)
 c. kono kare
 this he
 '*this he'

Modification by a restrictive relative clause is allowed as well.

- (18) Japanese (Sugamoto 1989:280)
 Nihongo ga hanas-eru kare wa fuijyuushi-nai
 Japanese NOM speak-can he TOP inconvenienced-NEG
 'He who can speak Japanese won't feel inconvenienced.'

However, like the Polish forms, the Japanese person forms are necessarily referential. They cannot, for example, be construed as bound variables, as shown by the ungrammaticality of (19b) as compared to (19a).

- (19) Japanese (Noguchi 1997:770)
 a. Mary ga [kanozyo ga tensai-da to] omotte-iru
 Mary NOM she NOM genius-COP COMP think-PRES
 'Mary_i thinks that she_i is a genius.'
 b. *Dono zyosei-mo [kanozyo ga tensai-da to] omotte-iru
 every woman also she NOM genius-COP COMP think-PRES
 'Every woman_i thinks that she_i is a genius.'

While this property is a pronominal one, it is outweighed by all the other nominal characteristics that the forms in question display. They are thus placed considerably lower on the pronominality scale in (10) than the English forms.

The expressions used to refer to discourse participants in Thai are even more noun-like than in Japanese. As in Japanese, they do not differ morphologically from nouns (both take no inflectional marking) and exhibit more or less the same modificational possibilities, which include modification by a numeral (20a), demonstrative (20b), quantifier-type expression (20c) and relative clause (20d).

- (20) Thai (Cooke 1968:10)
 a. khaw saam₂ khon₃
 3 three CLF
 'they three'
 b. phom₁ nii₂
 I: male this
 '*this I'
 c. raw₁ than₂ laaj₃
 1 all several
 'we all'
 d. raw₁ sy₂ pen₃ khon₄ ruaj₅
 1 who be rich persons
 'we who are rich'

Further, even more so than in Japanese, they do not constitute a closed class. The expressions regularly used to designate person include proper names, kin terms and various relational terms such as 'master', 'servant', 'individual crown of the head', etc. Cooke (1968) mentions twenty-seven specialized terms for the first person and twenty-two for the second person. (See ch. 6.) These, however, appear to be only a subset of the available possibilities. And significantly, the forms maintain much of their lexical meaning (similarly to *mummy* and *Johnny* in (1b)), and are highly diverse both stylistically and sociolinguistically. For example, the form *phom* 'you' is a general polite form, *taajthaa* is used only in highly deferential contexts, when speaking to a superior and *taajfàa'2la?ɔɔɔ'thulii'3phràbàad'4*, which literally means 'the one who is holding speaker under the dust of his foot', is employed only when addressing the king. While some or perhaps even most of the many forms that Thai has at its disposal to express person may indeed be fully nominal, the forms that are widely used such as *phom* are unlikely to emerge as such at least by virtue of their minimal semantic content.

In the light of the above, I will take the category of personal pronoun, in some sense of the term, to be universal. However, as I prefer to remain agnostic in regard to the nature of the morphological, syntactic and referential properties of personal pronouns, in order to avoid unnecessary confusion, I will refrain from using the term pronoun altogether. As stated earlier, in place of the term pronoun, the terms person form or person marker will be used.

1.3 The nature of this book

The existing similarities and differences in the properties of person markers, in person-marker inventories and in how person markers are deployed constitute a fascinating area of study. They have much to tell us about the human conceptual system and how it is organized. They provide important information about the relationship between the structure of language and the sociocultural and discourse conditions in which it is used. They offer significant insights into the processes of grammaticalization, that is the development, change and disappearance of grammatical categories and grammatical distinctions. And they constitute a source of crucial data for the determination of historical connections between languages, both genetic and areal. Moreover, the pervasiveness of person markers in language raises important questions in regard to their analysis within a theoretical model of grammar. The grammatical category of person manifests itself in both the nominal and the verbal domain and at various syntactic levels: the phrase, the clause, the sentence and even at the level of the text. The study of person thus takes us into each of these domains and levels of language structure and forces us to consider the nature of the relationships obtaining between them and how these should be analysed.

Needless to say, no one monograph devoted to the grammatical category of person can hope to do justice to all the above issues. Nonetheless, in the chapters that follow I hope to provide an overview of the most important concepts, controversies and analyses pertaining to the grammatical category of person, rich enough to constitute a good point of departure for the understanding of more detailed studies of specific issues or the carrying out of further in-depth research, be it synchronic or diachronic, descriptive or theoretical, on the category of person.

The structure of the book is as follows. Chapter 2 presents a typology of person markers from the point of view of the major factors underlying the existence of different person paradigms, namely morpho-phonological form, syntactic function and discourse function. In chapter 3 we turn to a consideration of how person paradigms are structured internally, that is to the nature of the person, number, gender and inclusivity oppositions that they express and how these are distributed in different types of paradigms, in the sense of chapter two. Chapter 4 explores the controversial topic of person agreement. The discussion is structured around the claim that there is no principled basis for distinguishing between anaphoric pronouns and person agreement markers and seeks to bring to light the commonality of the factors underlying the presence of anaphoric and grammatical agreement on different targets and with different controllers. Chapter 5 critically examines the unified account of the function of person markers developed within cognitive discourse analysis and captured in the relationship between relative discourse accessibility and morpho-phonological expression. We will consider to what extent relative cognitive accessibility may be invoked as the major factor underlying not only the inter-sentential but also the intra-sentential distribution of different types

of person forms and significantly, first- and second-person forms in addition to third-person ones. Chapter 6 looks at how person markers are utilized in the expression of social relations. The focus of the discussion is not on the specificities of the social relations that induce the use of special person forms, but on the person forms themselves, that is which forms are used and which are not, and which semantic oppositions are exploited in indicating social differences. Finally, chapter 7 seeks to place the preceding considerations in a diachronic context. It reviews the different sources of person markers and outlines the diachronic changes that they may undergo, including their grammaticalization pathways from independent markers to grammaticalized affixes and the factors that may underlie their eventual demise.

The discussion outlined above draws on language data from over 700 languages originating in the main from reference grammars and less frequently from theoretical descriptions of particular phenomena in individual languages or groups of languages. The subset of these languages, namely 402, are included in a computerized database which Dik Bakker and I have been developing for the last ten years. The relevant 402 languages will be referred to as the sample. Unless stated otherwise, the sample is the source of the observations pertaining to the distribution and frequency of particular properties and features of person markers that will be made. Originally the sample was compiled using the sampling methodology outlined in Rijkhoff and Bakker (1998), which aims for maximal genetic diversity. Since then, many languages have been added, even closely related ones. The sample is therefore merely a variety sample. In keeping with the nature of the sample, I have refrained from carrying out any serious statistics on frequencies and distributions and offer numerical data only as an illustration of clear tendencies and evident preferences. The 402 languages in the sample are listed by macro-area in Appendix 1.

2 The typology of person forms

Given the impoverished semantics of person markers and the fact that the range of syntactic and discourse functions that they fulfil cross-linguistically must essentially be the same, the major parameter responsible for the cross-linguistic variation in person markers is morpho-phonological form. Accordingly, this chapter will be concerned with the cross-linguistic formal realization of person markers and how this relates to their syntactic and discourse function.

We will begin our discussion with a review of the existing morpho-phonological types of person markers and their cross-linguistic distribution. Then we will proceed to examine the relationships between the morpho-phonological form of person markers and their syntactic function. This will involve a consideration of, on the one hand, the formal realization of different syntactic functions and, on the other, the alignment possibilities that the different morpho-phonological types of person markers enter into and how these relate to those found with lexical expressions. The last part of the chapter will deal briefly with the existence of person markers for special discourse functions.

2.1 Morpho-phonological form

The basic division of person markers in regard to morpho-phonological form is that between independent and dependent person markers. Other terms used for the independent forms are free, full, self/standing, cardinal, focal, strong, long and disjunctive. The dependent forms are also referred to as reduced, bound, defective, deficient or conjunctive. Typically what is meant by an independent/free/full, etc. person form is a person marker which constitutes a separate word and may take primary word stress, such as the English *I, me, you, she, they*. Dependent forms, by contrast, typically cannot be stressed (though some may receive contrastive stress), are often phonologically reduced relative to the independent forms, and either morphologically dependent on another element in the utterance or at least restricted in distribution relative to the independent forms. Since not all languages have dependent markers and the ones that do virtually always also have independent markers, we will consider the independent markers first.

2.1.1 Independent forms

Above I characterized an independent person marker as a marker which is a separate word and may take primary word stress. The notion of independence invoked is thus that of morphological and prosodic independence. Sometimes, however, what is meant by independent person marker is a syntactically independent form, that is a form capable of constituting a whole utterance all by itself, for instance as an answer to a question, as in (1b), or as an elliptical question, as in (2b).

- (1) a. Who are they going to ask?
b. Me./Her./Us.
- (2) a. He said that he would do it?
b. (Who) him?

Needless to say, syntactically independent person forms are also morphologically and prosodically independent, but the converse is not necessarily the case. For instance, in English the so-called possessive determiners, *my, your, our, their* are considered to be separate words which may be stressed but they cannot be used as utterances in their own right; when unaccompanied by a head noun the corresponding forms *mine, yours, ours*, etc. are employed (but see below). Thus (3b) does not constitute a possible response to (3a), only (3c) does.

- (3) a. Who are we going to invite, your mother or my mother?
b. *My.
c. Mine.

Nor may the English subject person forms *I, he, she*, etc., in contrast to the object forms, be used in complete isolation, as evidenced by the unacceptability of (4b) as compared to (4c).¹

- (4) a. Who wrote that?
b. *I/*He/*We
c. I did./He did./We did.

Yet they, too, are normally viewed as independent person markers. It must be pointed out, that not only certain person markers but also other elements which are normally considered to be words may not be able to constitute a separate utterance. For instance, most adpositions cannot occur alone. The same holds for articles. In view of the above, the ability to occur in complete isolation must be seen as being a too restrictive criterion for independent status. Therefore I will use the term independent form in the morphological and prosodic sense of the term independent, not the syntactic.

There are several diagnostics which may be employed in determining whether a particular person marker is or is not a separate word (see, e.g., Zwicky 1985;

¹ A response to (4a) involving a lexical NP, though preferred with the auxiliary *did*, would be possible without the auxiliary.