Chapter 5

Inclusive and exclusive in free and bound person forms

Anna Siewierska
Lancaster University

Dik Bakker
University of Amsterdam

It is generally recognised that in languages which have both free and bound person markers the semantic distinctions encoded in the latter may be less elaborate than those encoded in the former. This is typically attributed to the process of grammaticalization which in addition to a formal dimension also has a semantic one involving some semantic bleaching. The present chapter examines to what extent free and bound person markers differ with respect to the encoding of the semantic domain of inclusivity and to what extent less elaborate semantic encoding correlates with reduction in morpho-phonological encoding. The investigation is carried out on the basis of a cross-linguistic sample of 394 languages.

Keywords: person form, paradigm, grammaticalization, clitic, affix

1. Introduction

It is generally recognized that in languages which have more than one paradigm of person markers the semantic distinctions encoded in each paradigm need not be the same.[1] Differences between person paradigms in the nature of the semantic distinctions expressed are seen to be particularly likely in languages, which have both free and bound person markers. This is due to the fact that the latter are typically assumed to derive from the former via the process of grammaticalization, which, in addition to a formal dimension involving loss of independent morphological status and phonological reduction, may also have a semantic dimension, reflected in decrease of semantic distinctions encoded (cf. Givón 1976; Lehmann 1982b: 233–41; Bresnan & Mchombo 1986,1987; Siewierska 1999).[2]

While incidental examples documenting differences in paradigmatic structure in line with the above abound in the literature, there do not appear to be any studies devoted specifically to this issue. The present chapter seeks to fill in this gap in regard to the semantic dimension of inclusivity. It presents a detailed comparison of the encoding of the distinctions involving the groupings of speaker and hearer in free person markers and their bound counterparts in a cross-linguistic sample of 394 languages. The free person markers considered are morphologically independent nominative or absolutive forms or, in the case of languages lacking such forms, the so called cardinal forms that may be used as single word answers to questions

UNCORRECTED PROOFS
© JOHN BENJAMINS PUBLISHING COMPANY

such as "Who is coming?" The bound markers include weak pronouns, clitics and affixes functioning as the arguments of intransitive (S) and transitive clauses (A and P) and the possessors in adnominal possessive constructions (Poss).

The chapter is organised as follows. Section two presents and illustrates the approach to the domain of inclusivity that we have adopted and the nature of the semantic distinctions that we have made within this domain. Section three provides the data on the encoding of the domain of inclusivity in the free and bound person forms among the languages in our sample concentrating in particular on the paradigmatic differences encountered. In section four we consider to what extent the differences in the encoding of inclusivity among reduced person forms are related to their degree of grammaticalization with respect to morpho-phonological form, i.e. in relation to the weak pronoun vs clitic vs affix distinction and in relation to Siewierska's (1999) tripartite typology of person agreement markers, i.e. pronominal, grammatical and syntactic. Finally, section five offers some concluding remarks.

2. The Inclusive–exclusive distinction

Our approach to the domain of inclusivity is strongly based on the illuminating typology of person paradigms elaborated by Cysouw (2003) which takes as its point of departure the eight person distinctions shown in (1).

(1)

Singular	Group	
1+2	minimal inclusive	
1+2+3	augmented inclusive	
1	1+3	exclusive
2	2+3	
3	3+3	

In place of the familiar singular/plural distinction, this paradigm reflects a distinction between single persons and groups of more than one person. The singular category subsumes markers which refer to exactly one person or object and the group category markers that refer to more than one person or object, provided they are not specified with respect to the number of participants (e.g. dual, trial, paucal) and do not include within their scope regular reference to singular entities (Cysouw 2003: 66–72). We see that the paradigm differentiates between each of the three persons in the singular, having a distinct marker for each of the three persons, 1, 2 and 3. In the group it makes a five-way distinction between the speaker- addressee dyad 1+2, a configuration of the speaker, addressee and at least one other 1+2+3, a grouping of the speaker and at least one other but not the addressee 1+3, a constellation of the addressee and at least one other but not the speaker 2+3 and a grouping of at least two third parties 3+3. Thus what is of relevance for the group categories is not so much the number of participants, but their nature. Of particular signifi-

UNCORRECTED PROOFS
© JOHN BENJAMINS PUBLISHING COMPANY
1st proofs

cance in the context of the present discussion is, of course, the three-way distinction within the group first-person complex. The traditional inclusive–exclusive distinction involves the separation of groups involving the speaker and addressee, i.e. 1+2 & 1+2+3 from those that involve the speaker and some other party but not the addressee 1+3. In the paradigm in (1), however, the traditional category of inclusive is further subdivided into 1+2 (minimal inclusive) and 1+2+3 (augmented inclusive).

Needless to say, the actual person paradigms found in languages are frequently considerably less differentiated than the paradigm in (1). As far as the first-person complex is concerned, some paradigms have no separate form for non-singular first person, i.e. they have no "we" as distinct from "I". This is the case in Mura Piraha (2), a language isolate of Brazil, which according to Everett makes no distinction in regard to number of any type, not only in its person markers. The examples below are of the free forms.

(2) Mura Piraha (Everett 1987: 280–1)
1 ti
2 gíxai
3 hiapióxio

We will refer to paradigms such as that in (2) as exhibiting a *nowe* pattern.[3] Other paradigms have only one form for non-singular first person covering all three interpretations, 1+2, 1+2+3 and 1+3. Such paradigms, which we will call unified we (*unif we*) are familiar from English and most other European languages. For completeness an example is given in (3) from Dagbani, a Gur language spoken in northern Ghana.

(3) Dagbani (Olawsky 1999: 21)

Singular			Group		
			1+2		ti
			1+2+3		ti
1		ŋ	1+3		ti
2		a	2+2		yi
3	an	o	3+3	an	bE
3	inan	di	3+3	inan	di/ŋa

In yet other paradigms there is a special form for 1+2 & 1+2+3 but not for the exclusive 1+3. This pattern will be called only inclusive (*only incl*). It is illustrated in (4) from the Otomanguean language Chalcatongo Mixtec spoken in south-central Mexico, Oaxaca. The paradigm in (4) is of the free person forms but the corresponding clitics have the same paradigmatic structure.

UNCORRECTED PROOFS
© JOHN BENJAMINS PUBLISHING COMPANY
1st proofs

(4) Chalcatongo Mixtec (Macaulay 1996: 139)

		1+2	**žó?ó**
1	rú?ú	1+2+3	**žó?ó**
2	ró?ó		
3m	càà		
3f	nā?ā		
3animal	kɨtɨ		
3supernatural	í?a, íža		

In Chalcatongo Mixtec, as in other languages exhibiting the *only incl* pattern, the exclusive 1+3 combination is expressed by the first-person form accompanied by non-singular number marking as shown in (5b) and (5c) where *-rí* is the first-person clitic corresponding to *rú?ú*. Number is indicated in the language by a variety of morphological and syntactic means including the addition of the prefix *ka-* to the verb, as in (5b) and the use of the plural particle *xina?a* , as in (5c) or both.

(5) Chalcatongo Mixtec (Macaulay 1996: 139, 81,107)

a. *ni- žee-rí*
COMPL-eat-1
'I ate (it).'

b. *ka-satiu-rí*
PL-work-1
'We're working (exclusive of hearer).'

c. *kužaa-ri núndua xina ?á-ri*
live-1 Oaxaca PLURAL-1
'We (exclusive) will live in Oaxaca.'

In the case of the inclusive person forms additional number marking is also typically found as in (6) but appears not to be obligatory.

(6) *ni-ka-čaà-žó* *nùu bé?e*
COMPL-PL-arrive-1 face house
'We came to the house.'

A fourth type of subdivision of the first-person complex is the traditional inclusive–exclusive distinction (*incl/excl*) involving the grouping of of 1+2+3 & 1+2 as opposed to 1+3. An example is given in (7) from the Nilo-Saharan language So.

(7) So (Carlin 1993: 79)

Singular		Group	
		1+2	**inia**
		1+2+3	**inia**
1	aya	1+3	***isia***
2	bia	2+3	bitia
3	ica	3+3l	itia

UNCORRECTED PROOFS
© JOHN BENJAMINS PUBLISHING COMPANY
1st proofs

And finally there is the three way split within the first-person nonsingular where there are separate forms for the minimal inclusive, augmented inclusive and exclusive. Such a three way distinction, which we will term minimal augmented (*min/aug*) is found, for example, in the independent person markers of the Adamawa language Koh, spoken in Cameroon and Chad.

(8) Koh (Glidden 1985: 230)

Singular		Group	
1	mì	1+2	**ná**
		1+2+3	n*ári*
		1+3	bburu
2	mù	2+2	ì
3	ka	3+3	i

The above five subdivisions of the first-person complex have been captured by Cysouw (2003: 98) in the first-person complex hierarchy as shown in (9).

(9) The first-person complex hierarchy

no we > unif we > only incl > incl/excl > min/aug

This hierarchy represents the successive differentiation of the first-person complex from no separate form for *we* at all, through the presence of one undifferentiated form of *we*, then of a special form of *we* just for the inclusive category, followed by two different forms of *we* as in the traditional inclusive–exclusive distinction, and culminates in the three way distinction shown in the maximally differentiated paradigm in (1) above. This is depicted in Table 1 in the columns labelled (a) through (e), respectively. (The — means absence of a form.)

Given the above typology of the first-person complex, the dimension of inclusivity is richer than under the traditional inclusive–exclusive distinction which covers only the pattern in d), as it encompasses the possibilities of c) and e) as well.

The typology that we have adopted for this study differs from that of Cysouw's only in two minor respects, namely in the addition of two patterns both attested in Cysouw (2003: 91) and also discussed earlier by Greenberg (1988: 9).[4] The first of these is what we have called the minimal inclusive pattern (*min incl*) in which there is a separate from for the speaker-hearer dyad 1+2 and another form covering both 1+2+3 and 1+3.[5] This is illustrated in (10) on the basis of the independent pronouns

Table 1. The subdivisions of the first-person complex in Cysouw's (2001) hierarchy

	a	b	c	d	e
1+2	—	A	A	A	A
1+2+3					B
1+3			—	B	C

UNCORRECTED PROOFS
© JOHN BENJAMINS PUBLISHING COMPANY
1st proofs

in the Austronesian language Uma, spoken by around 20.000 people in Sulawesi.

(10) Uma (Martens 1988: 169)

Singular		Group	
	1+2	kita'	
	1+2+3	**kai'**	
1	aku'	1+3	**kai'**
2	iko	2+2	koi
3	hi'a	3+3	hira'

In the second additional pattern, which we will refer to as the augmented inclusive pattern (*aug incl*), there is a special form for 1+2+3 and another form covering 1+2 and 1+3 as exemplified in (11) from Bunuba a language of Western Australia.

(11) Bunuba (Rumsey 1999: 138)

Singular		Group	
1	ngayini	1+2	**ngiyirri**
		1+2+3	yaarri
		1+3	**ngiyirri**
2	nginji	2+2	yinggirri
3	niy	3+3	biyirri

The two readings of *ngiyirri* can be distinguished by the addition of number suffixes, the suffix *-way* for the 1+2 reading and the suffix *-yani* for the 1+3 reading.[6] Neither, however, is obligatory. The augmented inclusive pattern is de facto simply a dual vs plural distinction within the first person, as the 1+2+3 form is also used for any combination of 1+3 other than that involving only two individuals. We have included this pattern within the domain of inclusivity provided that the first-person dual is the only category specifically covering two individuals in the person paradigm. Otherwise we have recognised a distinct category of dual number.[7] Thus under our analysis the independent pronouns in a language such as Kapau, a Highland language of Papua New Guinea, are taken to display a straightforward dual/ plural contrast, as shown in (12), and not an augmented inclusive pattern, though exactly the same distinctions, within the first-person complex are made.

(12) Kapau (Oates & Oates 1968: 17, 45)

Sg		Du		Pl	
1	ni	1+2	yäl	1+2+3	nai
		1+3	yäl	1+3+3	nai
2	nti	2+2	qi	2+2+2	hai
3m	ago	3m+3	aqoä'u	3m+3+3	aqoä
3f	I	3f+3	isä'u	3f+3+3	i'yoä

In sum, our analysis of the distribution of the encoding of the inclusive–exclusive distinction in free and bound person markers will be based on the subdivisions of the first-person complex shown in Table 2.

UNCORRECTED PROOFS
© JOHN BENJAMINS PUBLISHING COMPANY
1st proofs

Table 2. The subdivisions of the first-person complex in our typology

	a	b	c	d	e	f	g
1+2	—	A	A	A	A	A	A
1+2+3				B	B		B
1+3			—		A	B	C

It must be pointed out that neither of the two additional patterns that we have included in our typology, i.e. pattern d) nor e) fits neatly into Cysouw's first-person complex hierarchy, as this hierarchy is conceived of as consisting of a series of interdependent choices such that a split in the inclusive category occurs only after the emergence of a distinct exclusive category. This is not the case in either the *min incl* pattern (d) or the *aug incl* pattern (e) as there is a special form for 1+2 and 1+2+3 respectively but no special form for the exclusive. In pattern d) the exclusive falls in with the augmented inclusive and in pattern e) with the minimal inclusive. Our typology reflects the number of special forms used for the expression of the groupings of 1+2, 1+2+3 and 1+3 without overlap with the 1 person singular form. It may be viewed as defining a hierarchy but only in the loose sense of the term. It does not constitute a typological hierarchy. Therefore our typology is expressed in (13) as an ordered grouping not in hierarchical form.

(13)	0 forms	nowe	(a)
	1 form	unified we	(b)
		only inclusive	(c)
	2 forms	minimal inclusive	(d)
		augmented inclusive	(e)
		inclusive/exclusive	(f)
	3 forms	minimal augmented	(g)

3. The data

The data for our study of the relationship between the encoding of the inclusive–exclusive distinction in free and bound person markers are drawn from a data base on person markers and grammatical relations currently consisting of 394 languages. The languages together with their areal distribution are presented in the appendix.[8] We took a fairly liberal interpretation of what constitutes a person marker which subsumes the celebrated noun-like expressions found in Thai, Vietnamese and other South-East Asian languages. Accordingly all but one of the languages in the sample (Mbay, a Nilo-Saharan language of southern Chad) emerged as having free person markers. One other language, Acoma, does have free person markers,

UNCORRECTED PROOFS
© JOHN BENJAMINS PUBLISHING COMPANY
1st proofs

but not for argument functions. As for bound person markers, we took into account affixes, clitics and also so-called weak pronouns.[9] The last of these are not actually either phonologically nor morphologically bound to a constituent. However, as they are phonologically reduced forms and differ from free forms in syntactic distribution (Bresnan 2001), we decided to group them together with clitics and affixes.[10] A case in point is that of what Sohn (1985) calls the subjectives in the Melanesian language Woleaian, which as shown in (14) function as person agreement markers.

(14) Woleaian (Sohn 1985: 150–51)
a. *Gaang i ta weri-Ø*
I 1SG not see-3SG
'I did not see it.'
b. Gaami gai lag!
you:PL 2PL go
'You(pl) go!'
c. *John ye weri-Ø Mary*
John 3SG see-3SG Mary
'John saw Mary.'

Only bound person markers in S or A or P or Poss function were taken into account.[11]

The encoding of the first-person complex among all the languages exhibiting free or bound person markers (for the first person) in terms of the typology presented in section two is shown in Table 3.

Before we proceed to consider these figures, a point of clarification is in order. Although, ideally one should apply the typology outlined in (13), separately to paradigms exhibiting a singular/group opposition and those in which there is a further contrast involving higher numbers, i.e. the dual and/or trial or paucal, as in Cysouw

Table 3. Frequency of the encoding of the first-person complex in the five types of person marker in the sample

N=394	Indep pro	S marker	A marker	P marker	Poss marker
NoWe	8 (2.0%)	26 (8.7%)	25 (8.3%)	37 (15.6%)	21 (8.5%)
UnifWe	224 (57.0%)	166 (55.0%)	165 (55.0%)	119 (50.2%)	135 (54.7%)
OnlyIncl	8 (2.0%)	15 (5.0%)	15 (5.0%)	11 (4.6%)	12 (4.9%)
MinIncl	6 (1.5%)	4 (1.3%)	4 (1.3%)	3 (1.3%)	3 (1.2%)
AugIncl	4 (1.0%)	2 (0.7%)	2 (0.7%)	2 (0.8%)	1 (0.4%)
Incl/Excl	124 (31.5%)	76 (25.0%)	74 (24.6%)	54 (22.8%)	68 (27.5%)
SplInclExcl	4 (1.0%)	3 (1.0%)	3 (1.0%)	4 (1.7%)	2 (0.8%)
MinAug	15 (3.8%)	8 (2.7%)	8 (2.7%)	7 (3.0%)	5 (2.0%)
TOTAL	393	302	300	237	247

UNCORRECTED PROOFS
© JOHN BENJAMINS PUBLISHING COMPANY
1st proofs

(2003), we have not done so. Consequently some of the encoding patterns in Table 3 must be understood as being relevant to whatever non-singular oppositions the paradigm displays, be it just group, or group and dual or group and dual and trial etc. This applies essentially to the *unif we* pattern, the *incl/excl* pattern and the *min/aug* pattern. Significantly, we have included under the *min/aug* pattern the famous five "we" paradigms found among some of the non-Pama Nyungan languages of Australia (see e.g. McKay 1978).[12] Under the traditional analysis these paradigms are seen as having an inclusive/exclusive contrast in the dual and plural and an inclusive trial encompassing 1+2+3. Under Cysouw's (2001: 265) analysis the trial is treated as a special type of dual. Instead of a straightforward inclusive/exclusive opposition in the dual and plural, there is an opposition between an augmented inclusive vs exclusive, and the minimal inclusive 1+2 is undifferentiated. This is shown in (15) on the basis of Mangarayi.

(15) Mangarayi (Merlan 1982: 102)

Sg		Du (restricted)		Pl (group)	
		1+2	ŋi		
		1+2+3	ŋa-r	1+2+3+3	ŋa-ḷa
1	ŋaya	1+3	ŋi-r	1+3+3	ŋi-ḷa
2	ñaŋgi	2+2	nu-r	2+2+3	nu-ḷa

Paradigms which display other differences in the encoding of inclusivity across the non-singular categories have been collapsed in Table 3 in the row labelled split inclusive/exclusive (*split incl/excl*). These include paradigms which are described as displaying:

- an incl/excl distinction only in the plural, as in the Papuan languageYava illustrated on the basis of the actor prefixes in (16);

(16) Yava (Jones 1986: 42)

Sg		Du		Pl	
		1+2	**ririm-**	1+2+3	*wam-*
1	sy-	1+3	**ririm-**	1+3+3	<u>ream-</u>
2	n-	2+2	ip-	2+2+3	wap-
3m	p-	3+3	y-	3+3+3	w-
3f	m-				

- an inclusive–exclusive distinction only in the dual, as in the case of subject prefixes in Tanimbili, an Oceanic language of Utupua Island, shown in (17);

(17) Tanimbili (Tryon 1994b: 628)

Sg		Du		Pl	
		1+2	*si-*	1+2+3	**misu-**
1	nyi-	1+3	<u>me-</u>	1+3+3	**misu-**
2	nu-	2+2	mwa-	2+3+3	muku-
3	I-	3+3	ŋgi(li)-	3+3+3	ŋgu-

UNCORRECTED PROOFS
© JOHN BENJAMINS PUBLISHING COMPANY
1st proofs

- a dual vs plural contrast only within the inclusive category, as in the Ponapean free person markers illustrated in (18) ;

(18) Ponapean (Rehg 1981: 158)

Sg		Du		Pl	
		1+2	*kita*	1+2+3	kitail
1	ngehi	1+3	**kihit**	1+3+3	**kihit**
2	kowe koh	2+2	kumwa	2+2+3	kumwail
3	ih	2+3	ira	3+3+3	irail/ihr

- a dual/plural contrast only within the exclusive category as in the free (and also clitic) forms in Yagua (19);

(19) Yagua (Payne 1990: 28–9)

Sg		Du		Pl	
		1+2	**nayin**	1+2+3	**nayin**
1	ráy	1+3	*nááy*	1+3+3	núúy
2	jiy	2+2	saadá	2+2+3	jiryéy
3	nii	3+3	naadá	3+3+3	riy

Let us now have a look at the figures in Table 3. We see that on the whole the distribution of the patterns of encoding of the first-person complex across the five types of person markers is strikingly similar. The two patterns which are of least interest to us, as they fail to differentiate between the inclusion and exclusion of the hearer in any way, have quite distinctive distributions. The *nowe* pattern is infrequent overall and also clearly less frequent in free person markers than in bound ones. The *unif we* pattern is the dominant pattern overall and prevails in all five person markers. Moreover its distribution within each type of person marker, given its commonality, is about the same.

Turning to the five patterns that do distinguish between the inclusion and exclusion of the hearer in one way or another, by far the most common is the traditional *incl/excl* pattern. Depending on the person form in question, the straightforward *incl/excl* pattern accounts for between 77% and 67% of the encodings of inclusivity, with the figure of 77% being manifested in the free forms. Of the three inclusive patterns, the *only incl* is the most common and the *aug incl* the least.[13] Interestingly enough the *only incl* pattern is slightly more frequent in the bound person markers than in the free forms. Quite the opposite holds for the *aug incl* pattern. The *spl incl/*excl pattern is attested with all the markers, but again very infrequently. As for the *min/aug* pattern, perhaps somewhat surprisingly, among the free person markers, it is the second most common form of encoding inclusivity but for the *incl/excl* pattern. The difference in the frequency of the latter relative to the former is, however, around 8:1.

Having reviewed the overall distribution of the encoding of the first-person complex in the five types of person markers, let us now turn to a consideration of the differences in the encodings exhibited by the various types of bound person markers

UNCORRECTED PROOFS
© JOHN BENJAMINS PUBLISHING COMPANY
1st proofs

relative to the free forms within languages. Only languages with both free and (at least one type of) bound person markers will be considered. [14]

Of the 330 languages which have both free and bound person markers, 179 (54%) have no encoding of the inclusive/exclusive domain. In other words, all the relevant person markers in question are either of the *nowe* or *unifwe* type. Of the remaining 151 languages which do encode the inclusive/exclusive domain in some subset of their person markers, 74 utilize the same pattern of encoding of the inclusive/exclusive domain for the free person markers and whatever set of bound person markers they display. Unsurprisingly, the most frequent pattern of encoding of the domain of inclusivity which is manifested by all the person forms in these 74 languages (77%) is the *incl/excl* pattern. It is exhibited in 57 of the 74 languages. In 43 (60%) languages, the inclusive–exclusive distinction is encoded only once in the paradigm, i.e. the various paradigms have no higher numbers. It is encoded twice (plural and dual) in 9 languages and three times (plural, dual and trial or paucal) in 5 languages. All the other patterns of encoding are also attested throughout the person paradigms found in languages, but rather infrequently. The relevant figures are summarised in Table 4.

Moving on from similarities in the encoding of inclusivity to differences, there are 77 languages which exhibit differences in the encoding of inclusivity among the person markers that they have. The relevant differences may be grouped into two types: increases in the distinctions encoded and decreases. The increases may in turn be subdivided into introductions of the distinction and further elaborations. The decreases, on the other hand, may be subdivided into complete loss of the distinction and reduction. Needless to say, whether a particular difference constitutes a case of increase or decrease depends on which pattern one takes as one's point of departure. As we are interested in the encoding of inclusivity in bound forms as compared to free forms, our point of reference will be the encoding pattern found in the free forms. We will view as cases of introduction any difference involving free forms manifesting a *nowe* or *unifwe* pattern and bound forms displaying one of the other of the encoding patterns. A case of introduction is shown in (20) on the basis

Table 4. Uniform encoding of the inclusivity domain

Pattern of encoding of inclusive–exclusive domain	Number of languages
Incl/Excl once in paradigm	43
Incl/Excl twice in paradigm	9
Incl/Excl three times in paradigm	5
Only Incl	6
Split Incl/Excl	2
Aug Incl	1
Min Incl	3
Min Aug	5
Total	74

UNCORRECTED PROOFS
© JOHN BENJAMINS PUBLISHING COMPANY
1st proofs

of the Tibetan language Chepang, in which the free person markers have a *unifwe* pattern, i.e. they do not exhibit any encoding of inclusivity, while the person affixes have an *incl/excl* pattern in both the dual and plural.

(20) Chepang (Caughley 1982: 54–5)

Free		Person suffixes	
1	ŋa	1	-ŋa
2	naŋ	2	-naŋ
3	ʔowʔ	3	Ø
1+2 ;1+3	ŋici	1+2	-ŋə-cə
2+2	niŋji	1+3	-təyh-cə
3+3	ʔoʔnis	2+2	-naŋ-jə
		3+3	-ce
1+2+3 ;1+3+3	ŋi	1+2+3	-ŋ-sə
2+2+3	niŋ	1+3+3	-təyh-ʔi
3+3+3	ʔowʔiem	2+2+3	-naŋ-sə
		3+3+3	-ʔi/sə

As instances of complete loss we will, in turn, consider any difference involving *nowe* or *unifwe* in the bound forms and one of the other patterns in the free forms. An example of complete loss is illustrated in (21) from the Mongolic language Daur in which the free person markers display a simple *incl/excl* pattern while the S and A person suffixes a *nowe* pattern.

(21) Daur (Wu 1996: 21, 27)

Free		SA suffixes in non-past	
1sg	bi:	1sg	bəəi-bi
2sg	ši:	2sg	bəi-ši
3sg	in	3sg	bəi-sul
1+2(+3)	bed	1pl	bəi-ba:
1+3	ba:		
2pl	ta:	2pl	bəi-ta:
3pl	a:n	3pl	bəi-sul

The other differences between free and bound forms, i.e., those which we have classified as involving further elaboration, or alternatively reduction are listed in (22).

(22) a. when no higher numbers are involved
- only incl vs all others (but for nowe, uniwe)
- aug incl or min incl vs incl/excl, min aug
- incl/excl vs min aug

b. when higher numbers are involved
- incl/excl vs incl/excl du (tr pau) pl
- split incl/excl vs inc/excldu (tr pau) pl
- incl/excldu(trpau)pl vs min aug

UNCORRECTED PROOFS
© JOHN BENJAMINS PUBLISHING COMPANY
1st proofs

An example of a reduction involving a *min/aug* pattern in the free forms and an *incl/excl* one in the bound is given in (23) from Tiwi, a non-Pama-Nyungan language of Northern Australia, in which the *incl/excl* pattern is found in the P suffixes. (The S/A prefixes display a *min/aug* pattern, like the free ones.)

(23) Tiwi (Osborne 1974: 54, 27)

Free		P prefixes	
1	ŋia	1	məni-
2	ŋiŋta	2	məṇI-
3m	ŋara	3	Ø-
3f	ŋira		
1+2	mua	1+2	**mani-**
1+3	ŋawa	1+3	məwəni-
1+2+3	ŋaya	1+2+3	**mani-**
2+2	nua	2+2+3	mani-
3+3	wuta	3+3	wəni

An instance of elaboration involving free person markers displaying a *min incl* pattern and bound S/A markers with a *min aug* pattern is presented in (24) from the West Papuan language Hatam.

(24) Hatam (Reesink 1999: 40)

Free		SA prefixes	
1	da	1	di-
2	na	2	a-
3	no(k)	3	Ø/ni-
1+2	*sa*	1+2	*si-*
1+2+3	**nye**	1+2+3	**i(g)-**
1+3	**nye**	1+3	ni-
2+2	je	2+2	ji-
3+3	yo(k)	3+3	i(g)-

Taking the patterns of encoding in free person markers as our starting point, the number of languages exhibiting a decrease in the encoding of the inclusivity domain in some set of their bound person markers is nearly four times as large as that displaying increases, namely 71 vs 15. This is what we would expect given the assumption that bound markers originate from free ones and, in the process of grammaticalization, often undergo semantic bleaching.[15] Semantic distinctions, however, may also be lost or reduced, over time, in free person markers. If this happens after the development of bound person forms, the result may be that the latter may manifest a distinction no longer in evidence in the former. Accordingly, the existence of increases in the encoding of inclusivity in bound markers vis a vis free forms is not surprising. Nonetheless, since the potential loss of semantic distinctions in free forms is likely to be due to factors other than grammaticalization, we may expect the increases to be far less common than the decreases. And this is indeed so. There

UNCORRECTED PROOFS
© JOHN BENJAMINS PUBLISHING COMPANY
1st proofs

is, of course, also the possibility of an inclusive–exclusive distinction developing in the free forms after the emrgence of the bound. As discussed by various scholars (e.g. Jacobsen 1980; Nichols 1992 and Sakel (this volume)), the inclusive–exclusive distinction appears to be particularly amenable to borrowing. Although typically instances of borrowing of the inclusive–exclusive distinction are not confined to just free forms, this is precisely what has happened in the Khoekhoe languages (!Ora, Eini, Nama). According to Güldermann (forthcoming) these languages have borrowed from the !Ui-Taa-family (Southern Khoisan) the first-person nonsingular form *si* and in doing so innovated the inclusive–exclusive distinction in their free forms. As we see in (25), on the basis of Nama, the bound forms in contrast to the free exhibit no inclusive/exclusive opposition either in the dual or the plural.

(25) Nama (Hagman 1973: 83, 87)

Free		Clitics	
1sg m	tiíta	1 m	-ta
1sg f	tiíta	1 f	- ta
1+2m	saákxm	1+2; 1+3 m	-kxm
1+2f	saám	1+2; 1+3 f	-m/-ìm
1+2c	saám	1+2; 1+3 c	-m/-ìm
1+3m	siíkxm	1+2+3; 1+3+3 m	-ke
1+3f	siím	1+2+3; 1+3+3 f	-se
1+3c	sim	1+2+3;1+3+3 c	-tà
1+2+3m	saáke		
1+2+3f	saáse		
1+2+3c	saáke		
1+3+3m	siítà		
1+3+3f	siíse		
1+3+3c	siíke		

As for the distribution of the subtypes of increases vs decreases in the four kinds of bound person marker, this is presented in Table 5. We see that the dominance of decreases vs increases holds very clearly for each bound person marker, the difference between the two ranging from over three times in the case of S and A markers to nine times in the case of the Poss markers. The fact that Poss markers display a slightly higher decrease in the encoding of the inclusivity domain vis a vis the free personal pronouns may be in part due to the fact that they are more closely related to free possessive pronouns than to the personal ones. Another reason for the above is that some languages have defective paradigms of bound possessor markers with bound forms only for certain person and/or number combinations. For example in Lele (Frayzyngier 2001: 61), an East Chadic language spoken in Chad, Koh Lakka (Gliggen 1985), a Niger-Congo language of the Adamawa branch spoken on the boarder of Chad and Cameroon and the North Australian language of the Daly River family Maranungku (Tryon 1970–16) there are bound possessor forms only for singular possessors, while non-singular ones are expressed by free forms. In the

UNCORRECTED PROOFS
© JOHN BENJAMINS PUBLISHING COMPANY
1st proofs

Table 5. Increases and decreases in the encoding of inclusivity in bound person markers relative to free person markers

Type of difference N=77	Smrk N=68	Amrk N=69	Pmrk N=57	Possmrk N=58
Introduction	9 (13%)	9 (13%)	5 (8.8%)	2 (3.4%)
Elaboration	2 (2.9%)	2 (2.9%)	2 (3.5%)	2 (3.4%)
Sum Increases	11 (16.2%)	11 (15.9%)	7 (12.3%)	4 (6.9%)
Loss	23 (33.4%)	22 (31.9%)	26 (45.6%)	25 (43.1%)
Reduction	13 (11.1%)	9 (13.0%)	9 (15.8%)	11 (19.0%)
Sum Decreases	36 (52.9%)	37 (53.6%)	35 (61.4%)	36 (62.0%)
Total diff	47 (69.1%)	48 (69.6%)	42 (73.6%)	40 (70.0%)
				41

Papuan language Moutain Koiali (Garland & Garland 1975: 420) and the Brazilian language Trumai (Guirardello 1999), by contrast, there are bound forms only for the third-person singular and non-singular. Again free person markers are imployed for the other person/number combinations. And in Adzera, an Austronesian language spoken in the Morobe Province of Papua New Guinea, the bound possessor forms distinguish only person but not number. Consequently, they are used in conjunction with free forms, which, unlike, the bound display both a number contrast and an *incl/excl* pattern. The paradigms of the relevant Adzera free and bound forms are given in (26) and examples of the co-occurrence of the bound froms with the free in (27).

(26) Adzera (Holzknecht 1986: 96–8)

Free		Poss
1	dzi	-ŋ?-gan?
2	u/agu	-m-gam
3	araŋan	-n-gan
1+2 (+3)	aga/agai	
1+3	agal	
2+2	agam	
3+3	rib	

(27) a. *dzi gudzu-ŋ?-gaŋ?*
1SG head-1-POSS
'my head.'

b. *agi gudzu-ŋ?-gaŋ?*
1INCL head-1-POSS
'our (inclusive) heads

c. *aga gudzu-ŋ?gaŋ?*
1EXCl head-1-POSS
'our (exclusive) heads

UNCORRECTED PROOFS
© JOHN BENJAMINS PUBLISHING COMPANY
1st proofs

Among the decreases, the vast majority are instances of complete loss. This again holds for each of the bound person markers. Although about three quarters of the losses are relative to a straightforward *incl/excl* pattern, all the other possibilities of complete loss are also attested in our data. An example involving *min/aug* and *nowe* is presented in (28) from the previously mentioned Australian language Maranungku.

(28) Maranungku (Tryon 1970: 15–16)

Free		Poss
1	ngany	-ni
2	nina	-la/-le/-li
3m	nankuny	-na
3f	ngankuny	-nga
1+2	**ngangku**	
1+2+3	*kitya*	
1+3	ngatya	
2+2	nitya	
3+3	witya	

Given that bound markers for the S, A, P and Poss need not all develop at the same time, nor from the same source, nor grammaticalize at the same rate, it is difficult to have any strong expectations in regard to which markers should most likely exhibit a decrease in the encoding of inclusivity vis a vis the free person forms. The best contenders are indubitably markers of the S and A since they are typically seen as developing before P markers or Poss markers. They could therefore be expected to have undergone more desemanticization than the P markers. However, no strong correlations emerge from the data in Table 5. Among the 28 languages with both bound A and P markers in which at least one of the two displays a decrease in the encoding of inclusivity vis a vis the free pronoun, there are only two in which the SA marker shows less encoding of inclusivity than the P marker. The first of these is the Austronesian language Yapese. We see in (29), that the SA weak forms display reduced encoding of inclusivity as compared to the free forms; the former exhibit an *only incl* pattern (there is a separate form covering 1+2 & 1+2+3 but the 1+3 forms are homophonous with the 1), the latter an *incl/excl* pattern both in the dual and plural. The P markers, unlike, the SA markers, exhibit the same distinctions as the free forms.

(29) Yapese (Jenssen 1977: 132–42)

	Free	SA	P
1	gaeg	**gu**	-eeg
2	guur	mu	-eem
3	qiir	i/Ø	Ø
1+2	gadow	da	-dow
1+3	gamouw	**gu**	-mow

UNCORRECTED PROOFS
© JOHN BENJAMINS PUBLISHING COMPANY
1st proofs

2+2	gimeew	mu	-meew
3+3	yow	ra	-row
1+2+3	gadaed	da	-daed
1+3+3	gamaed	**gu**	-maed
2+2+3	gimeed	mu	-meed
3+3+3	yaed	ra	-raed

The second language is the Australian language Nyulnyul, the person markers of which are depicted in (30).

(30) Nyulnyul (McGregor 1996: 23, 4047)

	Free	SA		P
		Nonfut	Fut	
1	ngay	nga-	nga-	-ngay
2	juy	mi-	wa-/mi-	-juy/ji
3	kinyingk	I-	yu-	Ø
1+2	*yay*	ya-	ya-	*-yay*
1+2+3	**yarrad**	ya-	ya-	**-yarrad**
1+3	**yarrad**	ya-	ya-	**-yarrad**
2+2	kurr	ku-	wa-	-kurr
3+3	irr/yirr	i-	yu-	-yiir

According to McGregor (1999: 23), the free person markers of Nyulnyul formerly displayed a *min/aug* pattern of encoding, with the form *yadir* used for 1+2+3. However, this form appears to have been lost, as it was lacking from McGregor's corpus. Consequently, we have classified the free person markers as displaying a *min incl* pattern. This is also the pattern displayed by the P suffixes. The SA prefixes, on the other hand, have a *unif we* pattern.[16] As the paradigms in (29) and (30) illustrate, in both languages the P markers are phonologically closer to the free forms than the S and A markers.

The opposite situation to the above, that is more encoding of inclusivity in A markers than in P markers occurs more frequently in our sample; it obtains in six of the 28 languages. Three of these languages, however, display a *nowe* pattern, i.e. they simply have no encoding of the first person at all, or no encoding of just the first-person non-singular. Accordingly, the possibility of encoding inclusivity does not arise. In another language, Hanis Coos, the P markers are always fused with the A markers, which is what may underlie the lack of encoding of inclusivity in the P markers. The remaining two languages which have less marking of inclusivity in P forms than in A forms are: Guaraní (31) and Kwaza (32), an unclassified language of Southern Rondonia, Brazil.

UNCORRECTED PROOFS
© JOHN BENJAMINS PUBLISHING COMPANY
1st proofs

(31) Guaraní (Gregores & Suarez 1967: 131, 141)

	Free	A/S$_{(A)}$	P/S$_{(P)}$
1	še	a-	če-
2	ne	re-	ne
1+2	**yané**	**ya-**	**yane-**
1+2+3	**yané**	**ya-**	**yane-**
1+3	oré	ro-	**yane-**
2+2	peẽ	pe-	pene-
3	o-	i-/iy-	

(32) Kwaza (van der Voort 2000: 154, 163, 169)

	Free	SA	P
1	si	-da	-ta
2	çyi	-ça	-nĩ'nã
3	ĩ	-Ø	-Ø
1+2	**txana**	**-a**	**-eteja**
1+2+3	**txana**	**-a**	**-eteja**
1+3	*tsi'tsğɛ* (*si* +*tsɛ*)	*-aça*	**-eteja**
2+2	çyi'tsɛ	-ça(ça)	-eteleçwa
3+3	ĩ	-Ø	-Ø

In both languages the free and SA forms display an *incl/excl* pattern, the P markers a *unifwe* pattern. We have no explanation to offer for why there should be no marking of inclusivity in the P forms.

The situation in regard to increases in the encoding of inclusivity among S, A, P and Poss markers relative to the free forms is analogous to that of decreases. No clear statistical implications emerge from our data. Analogously to what we observed in the case of the decreases, the overwhelming majority of the increases involve introductions rather than elaborations of the marking of inclusivity. The vast majority of the introductions are relative to a *unif we* pattern in the free forms, as shown earlier on the basis of Chepang. An introduction involving the other possibility, a *nowe* pattern, is illustrated in (33) from the Tanoan language Kiowa in which the inclusive–exclusive distinction is made only in the agentive prefixes. As (33) illustrates the inclusive is homophonous with the second-person plural and the exclusive with the third-person inverse.

(33) Kiowa (Watters 1984: 100, 113, 115)

	Free	S prefixes
1	nɔ́	à-
2 ám	èm-	
3	Ø-	
3 inverse	è-	
1+2(+3)	bà-	
1+3	è-	

UNCORRECTED PROOFS
© JOHN BENJAMINS PUBLISHING COMPANY
1st proofs

2+2	bà-
3+3(hum)	á-

4. The encoding of inclusivity and grammaticalization of person markers

Although, as pointed out by Lehmann (1982a: 236), we should not expect there to be an 100% correlation among all the various parameters of grammaticalization, the assumption underlying grammaticalzation theory is that the semantic, morphological and syntactic changes will tend to coincide. This suggests that the less elaborate encoding of the inclusivity domain in some bound as opposed to free person markers discussed above may have correlates on other dimensions of grammaticalization. We will consider two of these, namely the actual morpho-phonological form of the bound person markers and their co-occurrence possibilities with corresponding NPs.[17]

The bound person markers that we have been considering may be grouped in terms of their decreasing morphological independence and reduction in phonological form as in (34), with degree of grammaticalization increasing from left to right.

(34) weak > clitic > affix

All things being equal, the more grammaticalized markers from the point of view of morpho-phonological form should also exhibit more evidence of decrease in the encoding of the inclusivity domain, as compared to their free counterparts than the less grammaticalized markers. In other words, there should be relatively more instances of loss of the marking of inclusivity and reductions in the distinctions made among the affixes than among the clitics or weak forms.

The distribution of the three types of bound markers among the languages in the sample is very unbalanced as the affixes are overwhelmingly dominant and the weak forms very poorly attested. This holds for all four types of person markers, the S, A, P and Poss. Nonetheless, despite the skewed distribution, the relationship between morpho-phonological form and degree of encoding of the inclusivity domain (relative to free person markers) is as expected. This is documented in Table 6 on the basis of the S markers.

The differences between the types of encoding of inclusivity in S markers as opposed to free person markers in weak forms, clitics and affixes are marginal. They do, however, go in the right direction. There are no instances of loss among the weak forms and loss among clitics (2.2%) is marginally less frequent than in affixes (7%).

Slightly more robust is the relationship between the encoding of inclusivity in bound forms relative to free forms and degree of grammaticalization when viewed in terms of the tripartite typology of bound person markers in (35).[18]

UNCORRECTED PROOFS
© JOHN BENJAMINS PUBLISHING COMPANY
1st proofs

Table 6. Differences in the encoding of inclusivity between free forms and Smarkers relative to morpho-phonological form

Type of difference in the encoding of inclusivity N=296	Weak N=5	Clitic N=45	Affix N=246
Loss	0 (0%)	1 (2.2%)	18 (7.3%)
Reduction	1 (20.0%)	1 (2.2%)	11 (4.5%)
Increase	0 (0	1 (2.2%)	10 (4.1%)
No difference	4 (80.0%)	42 (93.3%)	207 (84.1%)
			208

(35) pronominal agreement markers > ambiguous agreement markers > syntactic agreement markers

This t4ypology builds on the widely held position that there is no good basis for distinguishing between agreement and antecedent-anaphora relations, even when cross-clausal (see e.g. Givón 1976; Moravcsik 1978; Lehmann 1982a, b; Corbett 1991; Anderson 1992). It is based on the co-occurrence possibilities between bound person markers and their corresponding NPs. Pronominal agreement markers are markers that are in complementary distribution with their corresponding NPs, be it nominal or pronominal, in the same construction. One language which has such person markers is the Carib language Macushi; note the presence of an S person prefix *aa-* in (36b) as opposed to (36a).

(36) Macushi (Abbott 1991: 84)
a. *u-yonpa-kon João ko'mamî-'pî miarî*
1-relative John remain-PAST there
'Our relative John stayed there.'
b. *aa-ko'mamî-'pî asakîne wei kaisarî*
3-remain-PAST two day up:to
'He remained two days.'

Ambiguous agreement markers are markers which occur both in the presence and the absence of a corresponding NP in the same construction as in the case of the S/A prefixes in Gumawana (37), an Oceanic language spoken in the Milne Bay province of New Guinea.

(37) Gumawana (Olson 1992: 326)
a. *Kalitoni i-paisewa*
Kalitoni 3SG-work
'Kalitoni worked.'
b. *i-situ vada sinae-na*
3SG-enter house inside-3SG(INAL)
'He entered the inside of the house.'

UNCORRECTED PROOFS
© JOHN BENJAMINS PUBLISHING COMPANY
1st proofs

And syntactic agreement markers are markers that necessarily require the presence of a corresponding NP in the same construction. Such markers are very infrequent outside western Europe. They are illustrated in (38) on the basis of Dutch.

(38) Dutch
a. *Hans zie-**t** Kees elke dag.*
Hans see-2/3SG Kees every day
'Hans sees Kees every day.'
b. **(Hij) zie-**t** Kees elke dag.*
he see-2/3SG Kees every day
'He sees Kees every day.'

The increasing degree of grammaticalization from pronominal via ambiguous to syntactic markers is reflected in their growing reliance on the presence of a corresponding NP in the same construction, from inability to co-occur with such an NP to the inability to occur without such an NP. Given the above, we expect relatively more instances of loss and reductions of the encoding of inclusivity relative to free forms the further to the right we proceed on the hierarchy in (35). Accordingly, we should find more instances of loss in syntactic markers than in ambiguous ones, and more in the latter than in pronominal markers. That this is indeed so is suggested by the figures in Table 7 again pertaining to S markers and excluding languages with bound forms only for the third person.

As in the case of the weak–clitic–affix distinction, the figures are heavily skewed due to the fact that distribution of pronominal, ambiguous and syntactic agreement markers is far from uniform. The vast majority of the bound person markers are grammatical ones. Pronominal markers are not uncommon, especially in the case of P and Poss forms. Syntactic markers in turn are very rare and appear to be found only in S or A functions. As the figures in Table 7 reveal, pronominal markers are the least likely and syntactic markers the most likely to exhibit loss or reduction in the encoding of inclusivity.

Table 7. Differences in the encoding of inclusivity between free forms and Smarkers relative to the tripartite typology of agreement markers

Type of difference in the encoding of inclusivity N=296	Pronominal marker N=32	Ambiguous marker N=257	Syntactic marker N=7
Loss	0 (0%)	18 (7.0%)	2 (28.6%)
Reduction	3 (9.4%)	10 (3.9%)	0 (0%)
Increase	1 (3.1%)	10 (3.9%)	0 (0%)
No difference	28 (87.5%)	219 (85.2%)	5 (71.4%)
			6

UNCORRECTED PROOFS
© JOHN BENJAMINS PUBLISHING COMPANY
1st proofs

5. Concluding remarks

Our comparison of the encoding of inclusivity in free and bound person markers has revealed that there is only a small difference in the frequency with which this semantic dimension is encoded in free as opposed to bound person markers. Free person markers are the most likely (42%) and P markers (34%) the least likely to display some encoding of inclusivity. Nonetheless, this does not mean that the type of inclusivity distinctions that are encoded in free and bound forms are necessarily the same. We have seen that among the 330 languages in our sample which have both free and bound person markers (of some type), the same pattern of inclusivity encoding is manifested in only 74, i.e. 22%. In nearly a quarter of the relevant languages there are some differences in how inclusivity is encoded in the various person paradigms. The vast majority of these differences involve presence of the encoding of inclusivity in free as compared to bound forms, though the reverse is also found. More subtle differences, involving changes in the type of inclusivity encoded are attested as well, such as that between minimal augmented and inclusive/exclusive, or inclusive/exclusive and only inclusive etc.

The loss or reduction in the encoding of inclusivity in bound person markers as compared to free ones correlates with two other facets of the grammaticalization of bound person markers, namely changes in their morpho-phonological form and their co-occurrence possibilities with corresponding NPs. The correlations are not as strong as the parallel path hypothesis would lead one to expect, even assuming Lehmann's (1982b: 236) caevet, mentioned at the beginning of section 4. There are at least two reasons for this. The first is the unbalanced distribution of affixes as compared to weak forms and clitics, on the one hand, and ambiguous as compared to both pronominal and syntactic markers on the other. The second is, the potential lack of a relationship between bound and free forms in some of the languages considered. Although, bound forms typically derive from free forms, the free forms in question are not necessarily those currently attested. The above notwithstanding, the data do suggest that loss and reduction of the encoding of inclusivity are more common among affixes than among weak or clitic forms, and more common among syntactic and ambiguous markers than among pronominal ones. Thus, at least as far as the encoding of inclusivity is concerned, the semantic parameter of grammaticalization is in tune with the morpho-phonological and syntactic. To what extent this holds also for other semantic features encoded together with person, such as number, gender, politeness etc. remains to be determined.

Abbreviations

1	first person
2	second person
3	third person
an	animate

UNCORRECTED PROOFS
© JOHN BENJAMINS PUBLISHING COMPANY
1st proofs

$AS_{(A)}$	actor
c	common gender
compl	completive
du	dual
excl	exclusive
f	feminine gender
hum	human
inal	inalienable
inan	inanimate
incl	inclusive
m	masculine gender
min/aug	minimal augmented
pau	paucal
pl	plural
poss	possessed
$PS_{(P)}$	undergoer
sg	singular
tr	trial
unif	unified

Appendix (N=394)

Africa (N=76)
Amharic, Ani, Arabic (Egyptian), Babungo, Bagirmi, Bambara, Bari, Beja, Berta, Bilin, Burunge, Coptic, Dagare, Diola-Fogny, Dizi, Dogon, Dongolese Nubian, Doyayo, Ewe, Fula, Fur, Geez, Grebo, Gude, Hamar, Hausa, Hebrew, Igbo, Iraqw, Kanuri, Katla, Kera, Kisi, Koh (Lakka), Kolokuma (Ijo), Koma, Kongo, Koromfe, Koyra Chiini, Kreol (Mauritian), Krongo, Kuku, Kunama, Lango, Lele, Luvale, Maale, Maba, Mbay, Mende, Mesalit, Mumuye, Mupun, Murle, Nama, Nandi, Ndonga, Ngiti, Nkore Kiga, Noon, Nupe, Oromo (Harar), Pari, Sandawe, Sango, So, Songhay (Koyraboro), Supyire, Swahili, Tamazight (Ayt Ndhi), Turkana, !Xu, Yaoure, Yoruba, Zande, Zulu

Eurasia (65)
Abkhaz, Ainu, Akkadian, Albanian, Armenian (Eastern), Basque, Bawm, Brahui, Burushaski, Byansi, Chepang, Chinese (Mandarin), Chukchi, Crimean Tatar, Dagur, Dong, Dutch, English, Evenki, Finnish, French, Garo, Georgian, German, Gilyak, Greek (Modern), Hindi, Hittite, Hungarian, Hunzib, Ingush, Irish, Italian, Japanese, Ju chen, Kannada, Kashmiri, Ket, Khalkha, Khasi, Korean, Kurdish (Central), Ladakhi, Lahu, Lak, Latvian, Lepcha, Lezgian, Limbu, Lushai, Meithei, Mundari, Nenets, Ossetic, Persian, Polish, Remo, Russian, Sema, Spanish, Sumerian, Turkish, Udihe, Welsh, Yukaghir (Kolyma)

South-East Asia (31)
Acehnese, Atayal, Burmese, Chrau, Hmong Njua, Indonesian, Kapampangan, Karo Batak, Kayah Li, Khmer, Khmu, Konjo, Larike, Makian (West), Mlabri (Minor), Muna, Paiwan, Rawang, Sahu, Savu, Semelai, Sundanese, Taba, Tagalog, Temiar, Thai, Tidore, Tsou, Tukang-Besi, Uma, Vietnamese

UNCORRECTED PROOFS
© JOHN BENJAMINS PUBLISHING COMPANY
1st proofs

Oceania (17)
Anejom, Chamorro, Dehu, Fijian, Kiribatese, Kusaiean, Lavukaleve, Malagasy, Maori, Mono Alu, Paamese, Palauan, Rapanui, Samoan, Tinrin, Ura, Yapese

New Guinea (48)
Abun, Adzera, Alamblak, Amele, Anem, Asmat, Au, Awtuw, Barai, Bukiyip, Daga, Dani (Lower Grand V), Ekari, Gapun, Gumawana, Hatam, Hua, Imonda, Kaliai Kove, Kapau, Kewa, Kilivila, Kobon, Koiali(Mountain), Koiari, Maisin, Marind, Maybrat, Nakanai, Nasioi, Salt (Yui), Selepet, Sentani, Suena, Tauya, Tawala, Tehit, Tigak, Tolai, Una, Vanimo, Wambon, Wanuma, Waskia, Yava, Yeletnye, Yessan-Mayo, Yimas

Australia (29)
Arabana, Bandjalang, Broken, Cape York Creole, Gooniyandi, Guugu Yimidhirr, Kalkatungu, Kayardild, Malakmalak, Mangarayi, Maranungku, Martuthunira, Maung, Ngalakan, Ngankikurungkurr, Ngiyambaa, Nunggubuyu, Nyulnyul, Panyjima, Pitjantjatjara, Tiwi, Ungarinjin, Uradhi, Wambaya, Wardaman, Yidin, Yukulta, Yulparija, Yuwaalaraay

North America (54)
Achumawi, Acoma, Atakapa, Cahuilla, Chumash Barbareno, Comanche, Comox, Cora, Cree (Plains), Greenlandic (West), Haida, Halkomelem, Hanis Coos, Jamul Tiipay, Karok, Kiowa, Koasati, Kutenai, Lakota, Makah, Maricopa, Mohawk, Mountain Maidu, Navajo, Nez Perce, Nootka, Oneida, Passamaquoddy, Quileute, Salinan, Seri, Slave, Southeastern Pomo, Squamish, Souther Sierra Miwok, Takelma, Tepehuan (Northern), Tlingit, Tonkawa, Tsimshian (Coast), Tunica, Umpqua (Lower), Wappo, Wasco-Wishram, Washo, Wichita, Wikchamni, Wintun, Yaqui, Yuchi, Yupik, Yurok, Zoque (Copainala), Zuni

Meso-America (14)
Chalcatongo Mixtec, Chinantec Lealao, Chocho, Copala-Trique, Jakaltek, Jicaque, Otomi (Mezquital), Pipil, Sierra Popoluca, Tarascan, Tetelcingo Nahuatl, Totonac (Misantla), Tzutujil, Zapotec San Lucas

South America (60)
Abipon, Amuesha, Apurina, Araona, Arawak (Lokono Dian), Awa Pit, Aymara, Barasano, Bororo, Bribri, Campa (Axininca), Candoshi, Canela Kraho, Capanahua, Carib, Cavinena, Cayuvava, Chacobo, Cubeo, Epena Pedee, Guaraní, Guaymi, Hixkaryana, Ika, Iquito, Jaqaru, Karitiana, Kawesqar, Kwaza, Makuchi, Mapuche, Marubo, Mataco, Miskito, Nadeb, Nambikuara, Ndyuka, Palikur, Paumari, Pech, Piraha, Quechua Imbabura, Rama, Retuara, Sanuma, Saramaccan, Selknam, Shipibo Konibo, Teribe, Tiriyo, Trumai, Urubu Kaapor, Waorani, Warao, Warekena, Wari, Waura, Witoto (Muinan), Xokleng, Yagua

Notes

1. We will use the term 'person marker' rather than 'pronoun' as not all bound person forms are pronominal, i.e. they cannot be used referentially.

2. This characterization of grammaticalization is, of course, a simplification. For a more detailed analysis of the parameters of grammaticalization see Lehmann (1982a: 234–41; 1982b); Heine & Reh (1984: 16–46), Bybee et al (1994: 19) and Croft (2000: 157–65).

3. The labels for the various patterns of encoding of the first-person complex that we are using are either directly taken over from Cysouw (2003) or are elaborations of his labels.

UNCORRECTED PROOFS
© JOHN BENJAMINS PUBLISHING COMPANY
1st proofs

4. Our minimal inclusive pattern corresponds to Cysouw's type P(f) and our augumented inclusive pattern to his type P(g). Both patterns were attested very poorly in his sample but are somewhat more common in ours and therefore have been included.

5. The only minimal inclusive pattern is referred to by Greenberg (1988: 9) as the Assinibone-type, but Assinibone can only be regarded as displaying the only minimal inclusive pattern if the separate number suffixes are taken into account. The paradigm of the actual person prefixes is of the straightforward unified we type.

6. The same suffixes can also be added to the second and third-person non-singular to achieve a dual vs plural reading.

7. This is also the position taken by Corbett (2000) and Cysouw (2003).

8. As we are interested in how the inclusive–exclusive distinction is encoded in free and bound forms, the sample is skewed in favour of languages which manifest the inclusive–exclusive distinction. Therefore the overall frequency of the encoding of this distinction is about 10% higher than what is suggested by Nichols (1992) or Bickel and Nichols (this volume).

9. We took as definitive of clitc as opposed to affixal status of a form its ability to attach to a variable host.

10. It is often difficult to determine whether a particular person marker is a clitic, e.g. a proclitic to the VP or a weak pronoun.

11. In the case of Poss markers, somewhat arbitrarily, we took into account only markers bound to the possessed, not to a classifier or an adposition.

12. These paradigms are often referred to as the unit augmented type. They are considered to be extensions of the minimal augmented type with the 1+2 category being treated as a "singular". For details see McKay (1978) and McGregor (1996b).

13. Crevels & Muysken (this volume) note that the *only incl* pattern is rather common among the languages of central-western South America. It occurs in just over 20% of the languages in the area which exhibit some marking of inclusivity.

14. We are not assuming that the current free and bound person markers in a language are necessarily related. The latter may have developed from forms which are no longer in existence. It must also be noted that the free forms may in fact be based on the bound

15. It is worth mentioning in this context La Polla's (this volume) findings with respect to the distribution of the encoding of inclusivity among the Tibeto-Burman languages. Of the 41 languages which encode the distinction among free forms only the Kiranti languages and some Chin languages also encode the distinction among their bound forms.

16. In Nyulyul the non-singular SA prefixes occur with additional plural markers which may be separted from the person prefixes by tense or mood prefixes. The *ya-* prefix sometimes occurs without a plural marker in which case it denotes 1+2. The 1+2 reading, however, is also available when there is an additional plural marker. Thus it appears that the ya-prefix, definitely is a unified we one.

17. That semantic reduction should be accompanied by phonetic reduction is particularly stressed by Bybee et. al (1994: 19).

18. This typology builds on the distinction between anaphoric and syntactic agreement discussed by Lehmann (1982a: 219) and the distinction between anaphoric and grammatical agreement elaborated by Bresnan & Mchombo (1986, 1987).

UNCORRECTED PROOFS
© JOHN BENJAMINS PUBLISHING COMPANY
1st proofs

References

Abbott, M. 1991. Macushi. In D.C. Derbyshire & G.K. Pullum (eds), *Handbook of Amazonian languages, Volume 3,* 23–160. Berlin: Mouton de Gruyter.

Andersen, S.R. 1992 *Amorphous morphology*. Cambridge: Cambridge University Press.

Bresnan, J. & Mchombo, S.A. 1986. Grammatical and anaphoric agreement. In A. Farley, P. Farley & K. McCullough (eds), *Proceedings of the Chicago Linguistic Society 22. Papers from the parasession on pragmatics and grammatical theory,* 741–82. Chicago IL: The University of Chicago Press.

Bresnan, J. & Mchombo, S.A. 1987. Topic, pronoun and agreement in Chichewa. *Language* 63: 741–82.

Bresnan, J. 2001. The emergence of the unmarked pronoun. In G. Legendre, J. Grimshaw & S. Vikner (eds), *Optimality-theoretic syntax,* 113–42. Cambridge MA: MIT Press.

Bybee, J., R. Perkins & Pagliuca, W.. 1994. *The evolution of grammar*. Chicago IL: The University of Chicago Press.

Carlin, E. 1993. *The So language*. Köln: Institut für Afrikanistik.

Caughley, R.C. 1982. *The syntax and morphology of the verb in Chepang*. Canberra: Australian National University [*Pacific Linguistics* B 84].

Corbett, G. 1991. *Gender*. Cambridge: Cambridge University Press.

Corbett, G. 2000. *Number*. Cambridge: Cambridge University Press.

Croft, W. 2000. *Explaining language change*. London: Longman.

Cysouw, M. 2003. *The paradigmatic structure of person marking*. Oxford: Oxford University Press.

Everett, D.L. 1986. Piraha. In D.C. Derbyshire & G.K. Pullum (eds), *Handbook of Amazonian languages, Volume 1,* 200–325. Berlin: Mouton de Gruyter.

Frajzyngier, Z. 2001. *A grammar of Lele*. Stanford: Stanford Monographs in African Languages.

Garland, R. & Garland, S. 1975. A grammar sketch of Mountain Koiali. In T.E. Dutton (ed.), *Studies in languages of Central and South-East Papua* 413–70. Canberra: Australian National University [*Pacific Linguistics* C 29].

Givón, T. 1976. Topic, pronoun and grammatical agreement. In C. Li (ed.), *Subject and topic,* 151–88. New York: Academic Press.

Glidden, S.H. 1985. The Koh verbal system. *Work Papers of the Summer Institute of Linguistics, University of North Dakota Session* 29: 223–82.

Greenberg, J. 1988. The first person dual as an ambiguous category. *Studies in Language* 12: 1–18.

Gregores, E. & Suarez, J.A. 1967. *A description of colloquial Guaraní*. The Hague: Mouton.

Guirardello, R. 1999. A reference grammar of Trumai. Ph.D. dissertation, Rice University, Texas.

Güldemann, T. Forthcoming. Die Entlehnung pronominaler Elemente des Khoekhoe aus dem!ui-Taa. In T.Schumann, M. Reh, R. Kiessling & L. Gerhard (eds), *Linguistische Beitrage zum 14. Afrkanistentag Hamburg 2000*. Köln: Köppe.

Hagman, R.S. 1973. Nama Hottentot grammar. Ph.D. dissertation, Columbia University. (also 1977 Research Center for Language and Semiotic Studies. Bloomington: Indiana University.)

Heine, B. & Reh, M. 1984. *Grammaticalization and reanalysis in African languages*. Hamburg: Buske.

UNCORRECTED PROOFS
© JOHN BENJAMINS PUBLISHING COMPANY
1st proofs

Jacobsen, W.H., Jr. 1980. Inclusive/Exclusive: A diffused pronominal category in native western North America. In J. Kreiman & A.E. Ojeda (eds), *Papers from the parasession on pronouns and anaphora. Chicago Linguistic Society, April 18–19, 1980,* 326–406. Chicago IL: The University of Chicago Press.
Jensen, J.T. 1977. *Yapese reference grammar*. Honolulu HI: University Press of Hawaii.
Holzknecht, S. 1986. A morphology and grammar of Adzera (Amari Dialect), Morobe province, Papua New Guinea. *Papers in New Guinea Linguistics* 24: 77–166.
Jones, L.K. 1986. The question of ergativity in Yawa, a Papuan language. *Australian Journal of Linguistics* 6: 37–56.
Lehmann, C. 1982a. Universal and typological aspects of agreement. In H.-J. Seiler & J. Stachowiak (eds), *Apprehension. Das sprachliche Erfassen von Gegenständen. Volume 2,* 201–67. Tübingen: Narr.
Lehmann, Ch. 1982b. *Thoughts on grammaticalization: A programatic sketch*, *Volume I.* Köln: Institut für Sprachwissenschaft [*Arbeiten des Kölner Universalien-Projekts* 48].
Macauley, M. 1996. *A grammar of Chalcatongo Mixtec.* Berkeley: University of California Press [*University of California Publications in Linguistics* 127].
Martens, M. 1988. Notes on Uma verbs. *Papers in Western Austronesian Linguistics* 4: 167–237.
McGregor, W. 1996. *Nyulnyul*. München: Lincom Europa.
McKay, G. 1979. Pronominal person and number categories in Rembarrnga and Djeebana. *Oceanic Linguistics* 17: 27–37.
Merlan, F. 1982. *Mangarayi*. Amsterdam: North-Holland [*Lingua Descriptive Studies* 4].
Moravcsik, E.A. 1978. Agreement. In J. Greenberg (ed.), *Universals of human language, Volume 4,* 331–74. Stanford: Stanford University Press.
Nichols, J. 1992. *Linguistic diversity in space and time*. Chicago IL: The University of Chicago Press.
Oates, W. & Oates, L. 1968. *Kapau pedagogical grammar*. Canberra: The Australian National University [*Pacific Linguistics* C 10].
Olawsky, K.J. 1999. *Aspects of Dagbani grammar*. München: Lincom Europa.
Olson, C. 1992. Gumawana (Amphlett Islands, Papua New Guinea): Grammar sketch and texts". In M.D. Ross (ed.), *Papers in Austronesian linguistics 2,* 251–430. Canberra: Australian National University [*Pacific Linguistics* A 82].
Osborne, C.R. 1974. *The Tiwi language*. Canberra: Australian Institute of Aboriginal Studies.
Payne, D.L. 1990. *The pragmatics of word order: Typological dimensions of verb-initial languages*. Berlin: Mouton de Gruyter.
Reesink, G.P. 1999. *A grammar of Hatam*. Canberra: The Australian National University [*Pacific Linguistics* C 146].
Rehg, K.L. 1981. *Ponapean reference grammar*. Honolulu HI: The University Press of Hawaii.
Rumsey, A. 1996. On some relationships among person, number, and mode in Bunuba. In W. McGregor (ed.), *Studies in Kimberley languages in honour of Howard Coate,* 139–48. München: Lincom Europa.
Siewierska, A. 1999. From anaphoric pronoun to grammatical agreement marker: Why objects don't make it. *Folia Linguistica* 33(2): 225–51.
Sohn, H. 1975. *Woleaian reference grammar*. Honolulu HI: University Press of Hawaii.
Tryon, D.T. 1970. *An introduction to Maranungku*. Canberra: The Australian National University [*Pacific Linguistics* B 15].

UNCORRECTED PROOFS
© JOHN BENJAMINS PUBLISHING COMPANY
1st proofs

Tryon, D. T. 1994. Language contact and contact-induced language change in Eastern Outer Islands, Solomon Islands. In T. Dutton & D. T. Tryon (eds), *Language contact and change in the Austronesian world,* 611–47. Berlin: Mouton de Gruyter.
van der Voort, H. 2000. A grammar of Kwaza. Ph. D dissertation, University of Amsterdam.
Watkins, L. J. 1984. *A grammar of Kiowa*. Lincoln: University of Nebraska Press.
Wu, C. (Ujiyediin Chuluu). 1996. *Daur*. München: Lincom Europa.

UNCORRECTED PROOFS
© JOHN BENJAMINS PUBLISHING COMPANY