

Who is who?

On the division of person space in pronominal paradigms

Michael Cysouw
University of Nijmegen

1 Introduction

This paper reports on a cross-linguistic study into the paradigmatical structure of pronominal elements. Pronouns are not considered as individual items, but as bound into a paradigm. The structure of such a paradigm determines the precise value of the item (cf. a phonological system or a Wortfeld in lexicography). One of the goals is to develop the intuitive concept of ‘richness’ of a pronominal paradigm more explicitly.

2 Definitions and Methodology

- Pronominal paradigms are defined as a group of morphemes that are syntagmatically equivalent (i.e. they form a ‘paradigm’) and that have at least an opposition between ‘speaker’ and ‘addressee’.
- Pronominal paradigms are taken as the locus of investigation. I do not talk about whole languages, but about paradigms within a language. The cases that are counted in the present paper are all *individual pronominal paradigms*. One language can have (and usually has) more than one pronominal paradigm.
- Pronominal paradigms can both be morphologically independent as well as inflectionally bound. I do not restrict a priori on the morphological status of the paradigm for it to be included in the sample.
- The presented data come from a breadth-search, not from a strictly controlled sample of the world’s languages. The first objective of this study is to give an overview of the variation of the structure of pronominal paradigms. Some regions or families turned out to be more interesting than others for this purpose, leading to some genetic skewing in the data. However, the complete sample is still good enough to qualify as a typological sound database. Included are around 400 languages, geographically as well as genetically well dispersed over the world. Because of the limited time and space, I will not expand on the actual data, but only present the generalisations.
- For the data of this paper, only paradigms with singular and general non-singular forms are included. All paradigms that distinguish gender, honorifics or specialised number categories are put aside. Only a few paradigms with gender in third person are included as otherwise some uncommon paradigmatical structures would have been missed. However, the traditional category ‘first person inclusive dual’ is included, as it is not interpreted as a dual, but as a ‘minimal inclusive’. These restrictions leave eight referentially different categories as the subject of the present paper.

		1+2	<i>minimal inclusive</i>] <i>inclusive</i>]] <i>first person plural</i>]
		1+2+3	<i>augmented inclusive</i>		
<i>speaker</i>	1	1+3	<i>exclusive</i>		
<i>addressee</i>	2	2+3	<i>second person plural</i>		
<i>other</i>	3	3+3	<i>third person plural</i>		

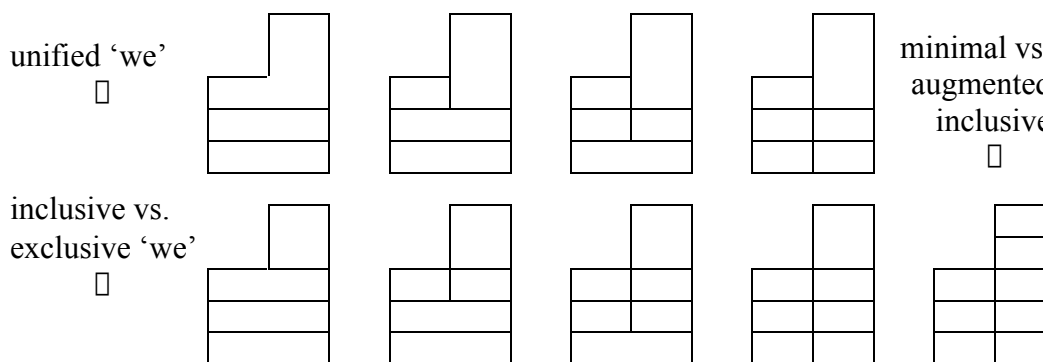
3 The First Person Complex and the Inclusive/Exclusive Opposition

A basic characteristic of the structure of a pronominal paradigm is the way in which the non-singular categories that include the speaker (1+2, 1+2+3 and 1+3) are marked. These three categories are called the ‘first person complex’. There are two main opposition that occur within these three categories: inclusive vs. exclusive’ and ‘minimal vs. augmented inclusive’. In the majority of cases, the second opposition is a subdivision of the first, although there are a few paradigms that do contradict this generalisation. For rest of the discussion, I will consider the four ‘exceptions’ as paradigms that have an inclusive/exclusive opposition.

		<i>Minimal Inclusive vs. Augmented Inclusive</i>		
		-	+	
<i>Augmented Inclusive vs. Exclusive</i>	-	144	4	148
	+	90	27	117
		234	31	265

Table 1: Implication ‘minimal vs. augmented \square inclusive vs. exclusive’ (Fisher’s exact $p = .000$; $\square = .315$)

There are nine different paradigmatical structures that occur frequently among the world’s languages. They are shown below. In the upper row, the paradigms are presented that have no inclusive/exclusive opposition. In the lower row, the common paradigms with an inclusive/exclusive opposition are shown.



The amount of paradigmatical variation is clearly less for paradigms that have an inclusive/exclusive opposition compared to the paradigms that have this opposition. This difference can most clearly be seen in the *amount of cases* with an *uncommon* paradigmatical structures.

		<i>Inclusive vs. Exclusive</i>		
		-	+	
<i>Common Paradigms</i>	4 (87 cases)	5 (97 cases)	9 (184 cases)	
<i>Uncommon Paradigms</i>	32 (57 cases)	21 (24 cases)	53 (81 cases)	
		36 (144 cases)	26 (121 cases)	62 (265 cases)

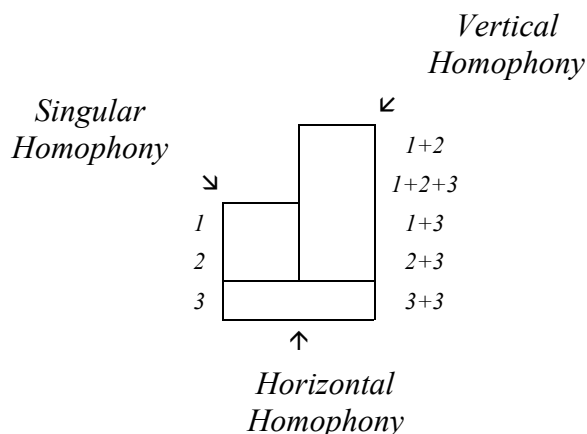
Table 2: Implication ‘Inclusive/exclusive opposition \square less cases of uncommon paradigms’ (Fisher’s Exact $p = .001$; $\square = -.241$)

4 The Analysis of Variation and the concept ‘Pure Person’

For the rest of this paper, I will formulate structural generalisation over all 265 cases of the 62 different paradigmatical structures. Three different characteristics are used to analyse the structure of a pronominal paradigm. The term ‘homophony’ is used in a purely descriptive meaning, relative to the present set of eight possible referential categories. If some of the 8 referential categories are marked by the same morpheme, this is called a ‘homophony’.

- **Vertical Homophony** is the phenomenon that the categories 2+3 or 3+3 (or both) are marked identically to one of the morphemes of the first person complex (e.g. the German suffix ‘...-en’);
- **Singular Homophony** is the phenomenon that the three singular categories 1, 2 and 3 are only marked by two different morphemes (e.g. the Dutch suffixes ‘...-Ø’ and ‘...-t’);
- **Horizontal Homophony** is the phenomenon that one morpheme marks both for a singular as well as for a non-singular category (e.g. the English pronoun ‘you’).

A few examples of these characteristics are presented in the picture below. There are many more possibilities, though, that are not shown.



The inclusive/exclusive opposition is correlated with vertical homophony: paradigms with an inclusive/exclusive opposition do hardly show vertical homophony. Within the group of 14 ‘exceptions’, by far the most common variants of vertical homophony are ‘inclusive with second person plural’ (5 cases) and ‘exclusive with third person plural’ (4 cases). These patterns can be referentially explained as, on the one hand, the *inclusion of the addressee* is relevant for both the inclusive and the second person plural; oppositely, the *exclusion of the addressee* is relevant for both the exclusive and the third person plural. The remaining 5 ‘exceptional’ cases show all a different kind of homophony.

		Inclusive vs. Exclusive		
		-	+	
Vertical Homophony	-	103	107	210
	+	41	14	55
		144	121	265

Table 3: Implication ‘inclusive/exclusive opposition \square less vertical homophony’ (Fisher’s exact $p = .001$; $\square = -.208$)

The inclusive/exclusive opposition is also correlated with singular homophony: paradigms with an inclusive/exclusive oppositions do not show singular homophony. This implication is even stronger as the former. Both correlations can be explained by the concept ‘pure person’ marking.

		<i>Inclusive/Exclusive</i>		
		–	+	
<i>Singular Homophony</i>	–	119	121	240
	+	25	0	25
		144	121	265

Table 4: Implication ‘inclusive/exclusive opposition’ \square no singular homophony’
(Fisher’s exact $p = .000$; $\square = -.296$)

‘Pure person’ marking refers to the fact that paradigms with an inclusive/exclusive opposition distinguish consequently all different referential possibilities that can be made with the two most important interactional categories: the speaker and the addressee. The reference to ‘speaker and others’ is distinguished from the reference to ‘addressee and others’, but also from reference to a group that includes ‘speaker *and* addressee’. In such a paradigm, the referential values are of such importance that it is not possible to mix up the reference elsewhere in the paradigm. Only when the inclusive/exclusive opposition is not present, other referential mixes are possible. Then it is possible to have vertical and/or singular homophony in the paradigm. Note that also the majority of the ‘exceptional’ cases with an inclusive/exclusive opposition *but with* a vertical homophony sort of follow the ‘pure person’ hypothesis as the inclusive is mixed with the second person plural and the exclusive with the third person plural. The reference to speaker and addressee is preferably not intermingled.

Finally, the inclusive/exclusive opposition is not correlated with horizontal homophony. This shows that although the *kind of participants* are important for paradigms with an inclusive/exclusive opposition, the *co-participants* besides the speaker and the addressee are of lesser importance.

		<i>Inclusive/Exclusive</i>		
		–	+	
<i>Horizontal Homophony</i>	–	69	65	134
	+	75	56	131
		144	121	265

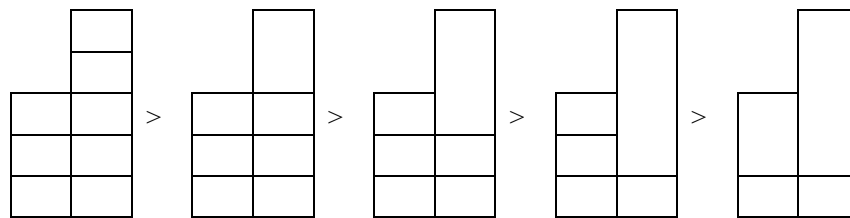
Table 5: No correlation between inclusive/exclusive opposition and horizontal homophony
(Fisher’s exact $p = .388$; $\square = -.058$)

5 Explicitness Hierarchy

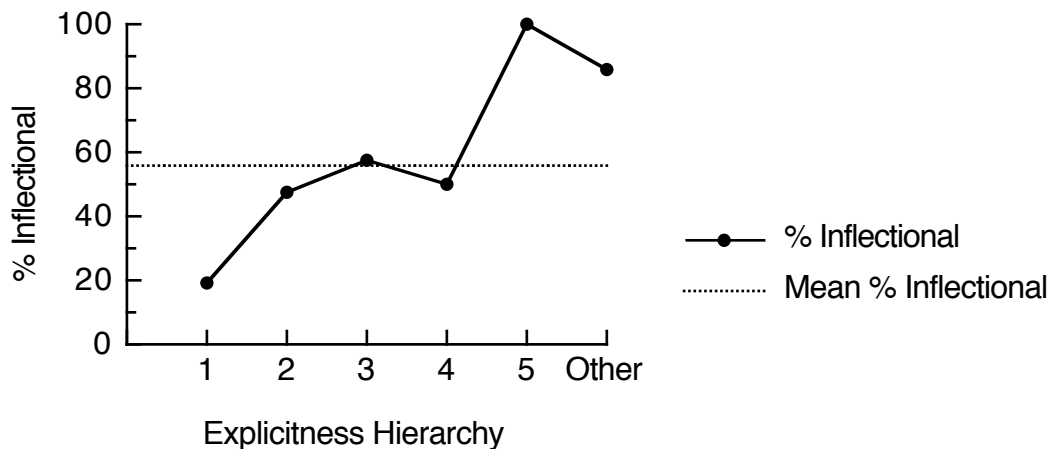
The various characteristics of a pronominal paradigms that were shown to be related actually form a hierarchy. High on the hierarchy are paradigms that maximally distinguish different referential categories by different morphemes. Low on the hierarchy are paradigms that barely distinguish the possible referential categories.

<i>Minimal vs. Augmented Inclusive</i>	+	-	-	-	-	+	-	-	+	+
<i>Inclusive vs. Exclusive</i>	+	+	-	-	-	-	+	-	+	-
<i>No Vertical Homophony</i>	+	+	+	-	-	+	-	+	-	-
<i>No Singular Homophony</i>	+	+	+	+	-	+	+	-	+	+
Amount of cases	26	78	99	20	21	3	12	4	1	1
	(92.1% of total)					(7.9% of total)				
<i>Independent</i>	21	41	42	10	0	3				
<i>Inflectional</i>	5	37	57	10	21	18				
% Inflectional	19.2	47.4	57.6	50.0	100	85.7				

This hierarchy can be thought of as showing more and more homophony from left to right, starting from the first person complex, but invading the rest of the paradigms sequentially. In the picture below, only a prototypical example of each stage is shown.



This Explicitness-hierarchy is correlated with morphological status. Highly explicit paradigms are more often independent, barely explicit paradigms are more often inflectional. The ‘exceptional’ cases are more often inflectional than independent. Note that the mean percentage of inflectional paradigms over all 265 cases is not 50 % but 55.8 % (148 inflectional versus 117 independent).



6 Conclusion and Prospects

The main finding of this cross-linguistic survey is the concept of ‘pure person’ marking. For ‘pure person’ marking, all reference to speaker and addressee (in different combinations) has to be separated in the linguistic marking. An important characteristic of ‘pure person’ marking is the opposition between inclusive and exclusive first person plural. Paradigms that distinguish between inclusive and exclusive first person plural show much less variation than other paradigms. Not only the amount of variation is less, also the kind of variation is much more structured. Paradigms with an inclusive/exclusive opposition (i.e. with ‘pure person’) do not allow for vertical and singular homophony, as these kind of morphemes mix up the reference to the basic speech act participants speaker and addressee.

The restrictions on the possible homophony can be captured in the ‘explicitness hierarchy’. This hierarchy is correlated with morphological status of the paradigms. Paradigms that are highly explicit are more commonly found to be morphologically independent; paradigms that are barely explicit are more commonly found to be inflectional. An interesting extra generalisation is that ‘exceptions’ to the formulated generalisations show an even higher percentage of inflectional cases than the ‘regular’ kinds of paradigms. Exceptional paradigmatic structures more easily arise in inflectional paradigms.

Part of an explanation for this difference in morphological status of the various paradigms can be found in the amount of *awareness* people have of a linguistic element. Independent words are like ‘things’ to a language user. They are highly structured by the difference in referential value, as this is important in daily usage of language. Oppositely, inflectional categories are often not very penetrable for the language user. In these paradigms, much more is possible.

Finally, a few prospects into other distinctions that seem to pattern with the presented generalisations. First, paradigms that distinguish number marking (dual, trial, paucal) seem to follow the explicitness hierarchy. Vertical homophony is not attested when there is an inclusive/exclusive opposition. Also, paradigms with an inclusive/exclusive opposition seem to be more often independent than the paradigms without this opposition. Second, gender marking is occasionally found in categories that include the speaker or the addressee. However, following the ‘pure person’ hypothesis, gender is not found in these categories when an inclusive/exclusive opposition is present in the paradigm (although gender can still be found in the third person in these paradigms). In a ‘pure person’ paradigm only person marking counts. No other categories are allowed to intermingle.

Drs. Michael Cysouw
General Linguistics and Dialectology (ATD)
University of Nijmegen
Postbus 9103
6500 HD Nijmegen

e-mail: m.cysouw@let.kun.nl