

On the genesis of personal pronouns: Some conceptual sources

BERND HEINE AND KYUNG-AN SONG*

University of Cologne and Chonnam National University

Abstract

Personal pronouns belong to the most conservative parts of grammar; that is, they are diachronically fairly stable, as suggested by the fact that in many cases they can be traced back etymologically to or even beyond the earliest stages of reconstructible language history. This is why so far not much progress has been made in unraveling the genesis of personal pronouns: the number of cases where it has been possible to determine their origin is severely limited. Still, some data do exist and allow for the reconstruction of some major lines of grammatical evolution. As we hope to show in this paper, this evolution is the product of a handful of crosslinguistically fairly uniform cognitive and communicative strategies.

Keywords

grammaticalization, identifier, personal pronoun, pluralization, shift in deixis, spatial deixis

1. Introduction

That it is possible to reconstruct human languages back to their earliest stages of evolution has been argued for by Heine and Kuteva (2007),

* Contact address: Bernd Heine, Department of Linguistics and African Studies, University of Köln, 50923 Köln, Germany. E-mail: bernd.heine@uni-koeln.de. Acknowledgements: We wish to express our gratitude to a number of colleagues who have been of help in writing this paper, in particular to Tania Kuteva, John Haiman, Heiko Narrog and Fritz Newmeyer. Furthermore, we have to thank an anonymous reviewer of the journal for highly constructive comments, which have all been extremely useful in revising the manuscript. Finally, our gratitude is also due to the Korean Ministry of Education, Science and Technology for generously having sponsored the research leading to this paper within its World Class University program (Grant No. R33-2008-000-10011-0).

using grammaticalization theory as a tool. But personal pronouns have hardly played any role in that work. In the present paper it is argued that it is nevertheless possible to reconstruct some of the cognitive forces that can be held responsible for the rise and development of personal pronouns.

The paper is organized as follows. The present section provides a brief outline of grammaticalization theory insofar as it relates to the subject matter discussed here. Section 2 is the main part of the paper, where we attempt to reconstruct salient conceptual sources of personal pronouns. While our concern in the paper is primarily with individual forms of personal deixis, we will deal in Section 3 with the question of how an entirely new system of personal pronouns may arise by looking at the example of a system that has evolved in the course of the last few centuries. In Section 4 finally we will highlight some general findings presented in the paper.

Unlike many other works that are concerned more generally with person markers (e.g. Forchheimer 1953; Cysouw 2003; Siewierska 2004), our interest here is mainly with (independent, or free) personal pronouns, which we define as words whose primary or only function it is to express distinctions of personal deixis; these have their own word prosody and resemble noun phrases in their positional possibilities (cf. Heath 2004: 1002). Languages are generally assumed to have personal pronouns.¹ We will treat personal pronouns as a subclass of person markers² (the latter of which include a wide range of other grammatical elements, like bound forms and agreement markers), whose main function is to distinguish speech-act participants. The boundaries between personal pronouns and other kinds of person markers are fluid³, and whether or not a given element conforms in every respect to our definition of personal pronouns is not a matter of great concern in this paper. In many typological works, third-person pronouns are distinguished from the other two categories of

-
1. At least two languages are reported to lack personal pronouns, namely the Keresan language Acoma of New Mexico and the Chapacura-Wanhan language Wari' of Brazil (Heath 2004: 999). We are ignoring here views based on specific theoretical assumptions that led the authors concerned to distinguish between different levels of analysis. For example, on the basis of syntactic evidence, Postal (1969) concludes that English does not have personal pronouns; rather, 'surface structure' forms such as *he*, *she*, or *we* are in their 'deep structure' articles similar to *the* and *a*.
 2. Person markers are linguistic elements that are shifters which are specialized for this function and are used for reference to speech act participants (Cysouw 2003: 5).
 3. Siewierska (2004: xv) rightly points out that "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". Unfortunately, there is so far no diachronic reconstruction of this cline based on crosslinguistic data—a task that can also not be tackled in the present paper.

personal deixis and there are good reasons to do so: For example, third person pronouns do not refer to speech participants and tend to differ in their morphosyntactic behavior from first- and second-person pronouns. We follow Siewierska (2004: 8), however, in maintaining that there is no clear advantage in excluding third-person pronouns from discussions of the category of person; as we hope to demonstrate below, this is also borne out by the study of their evolution.

Grammaticalization, which is the central theme of this paper, is defined as the development from lexical to grammatical forms and from grammatical to even more grammatical forms.⁴ Since the development of grammatical forms is not independent of the constructions to which they belong, the study of grammaticalization is in the same way concerned with constructions and with even larger discourse segments (see Traugott and Heine 1991a, 1991b; Heine et al. 1991; Hopper and Traugott 2003; Bybee et al. 1994; Lehmann 1995; Heine and Kuteva 2002 for details). In accordance with this definition, grammaticalization theory is concerned with the genesis and/or development of grammatical forms. Its primary goal is to describe how grammatical forms and constructions arise and develop through space and time and to explain why they are structured as they are.

In order to describe grammaticalization processes, a wide range of criteria have been proposed (see e.g. Lehmann 1995; Heine et al. 1991; Hopper and Traugott 2003; Bybee et al. 1994). Our model is based on the four parameters listed in (1). A number of alternative criteria have been proposed, such as syntacticization, morphologization, ‘obligatorification’, subjectification, etc. We argue that these fall under the four parameters in (1). Henceforth we will rely on these parameters, using them as a tool for identifying and describing instances of grammaticalization.

(1) Parameters of grammaticalization

- a. Extension, i.e. the rise of new grammatical meanings when linguistic expressions are extended to new contexts (context-induced reinterpretation),
- b. Desemanticization (or ‘semantic bleaching’), i.e. loss (or generalization) in meaning content,
- c. Decategorialization, i.e. loss of morphosyntactic properties characteristic of lexical or other less grammaticalized forms, and
- d. Erosion (‘phonetic reduction’), i.e. loss in phonetic substance.

4. For a fairly comprehensive list of definitions that have been proposed for grammaticalization, see Campbell and Janda (2001).

Each of these parameters concerns a different aspect of language structure or language use; (1a) is pragmatic in nature⁵, (1b) relates to semantics, (1c) to morphosyntax, and (1d) to phonetics. Except for (1a), these parameters all involve loss of properties. But the process cannot be reduced to one of structural ‘degeneration’: There are also gains: In the same way that linguistic items undergoing grammaticalization lose semantic, morphosyntactic and phonetic substance, they also gain properties characteristic of their uses in new contexts⁶—to the extent that in some cases their meaning and syntactic functions may show little resemblance to their original use.

The ordering of these parameters reflects the diachronic sequence in which they typically apply: Grammaticalization tends to start out with extension, which triggers desemanticization and subsequently decategorialization and erosion. Erosion is the last parameter to come in when grammaticalization takes place and in a number of the examples to be presented below it is not (or not yet) involved. Paradigm instances of grammaticalization involve all four parameters but, as we will see below, there are also cases where not all of the parameters play a role.

The question of what status should be assigned to personal pronouns in a taxonomy of grammatical categories has been a subject of controversy. We will assume that personal pronouns are functional categories rather than, say, a subclass of nouns or of determiners and in doing so we are relying on the parameters of grammaticalization proposed above: (a) compared to the lexical categories of nouns and verbs, personal pronouns have a schematic meaning that can be described fairly exhaustively in terms of a few elementary conceptual distinctions, being restricted in many cases to deixis and number (desemanticization); (b) personal pronouns have a more restricted categorial potential than lexical categories, frequently lacking e.g. the ability to take modifiers or inflectional and derivational morphology (decategorialization); and (c) personal pronouns are as a rule shorter than nouns and verbs (erosion).

-
5. For good reasons, an anonymous reviewer wonders whether this is really a pragmatic parameter. One should in fact be aware that there are also semantic factors that play a role in extension and that it is hard, if not impossible, to determine exactly to which extent each of the two is involved in a given case.
 6. The semantic gains have been described more recently as ‘procedural information’ (Nicolle in press).

2. Conceptual sources

In spite of all the diversity that is associated with both the crosslinguistic and language-internal structure of personal pronouns, we argue that their evolution is fairly regular: First, it is essentially unidirectional and, second, it involves only a handful of main conceptual sources, namely the ones listed in (2). In the remainder of this section we will be concerned with these sources and their contribution to the creation of new forms of personal deixis.

- (2) The main conceptual sources of personal pronouns
 - a. Nominal concepts
 - b. Spatial deixis
 - c. Identifiers
 - d. Pluralization
 - e. Shift in deixis.

2.1. *Nominal concepts*

Presumably the main way in which new forms of personal deixis are created is by transferring nominal expressions from the domain of referential objects, which in most cases are nouns or noun phrases referring to humans, to the domain of social deixis. That personal pronouns are the result of the grammaticalization of nouns has been argued by a number of authors since Humboldt (1830) (Blake 1934; Cysouw 2003: 13; Siewierska 2004). This is most obvious in the languages of Southeastern Asia: Languages like Thai, Burmese, Khmer, and Vietnamese have large repertoires of first- and second-person pronouns many of which have properties in common with nouns (Cooke 1968; Heath 2004: 1013), and John Haiman (p.c.) observes that most of the personal pronouns in Khmer “are in fact common nouns capable of modification and quantification”.

Within the class of nominal expressions, certain semantic subclasses are exploited time and again for grammaticalization. These subclasses include those listed in (3).

- (3) Domains of nouns recruited for the grammaticalization of personal pronouns
 - a. Honorific expressions used for high-ranking personalities
 - b. Terms for social status distinctions
 - c. Kin terms
 - d. Terms for professions

- e. Generic nouns for human beings, or based on distinctions of sex and age

Furthermore, there are nouns for ‘body’ that have given rise to markers of personal deixis in a number of languages. It would seem, however, that in such cases it was not a straightforward grammaticalization from noun to personal pronoun; rather, there was an intermediate stage of an identifier; hence we will discuss this process in Section 2.3. When nouns (or noun phrases) are grammaticalized to personal pronouns, this may take two different courses. The first—let us call it the *noun-to-pronoun channel*—is by far the most common one. The second course may be referred to loosely as the *noun-ellipsis channel*, whereby a nominal argument in a predication is deleted but there remains a third-person anaphor that bears witness to its earlier presence. With the disappearance of the nominal argument, the anaphoric argument assumes the function of a personal pronoun. We will return to this evolution in Section 2.5.

The noun-to-pronoun channel involves a straightforward process of grammaticalization from nominal to pronominal form, as sketched in (4).

- (4) Grammaticalization: The noun-to-pronoun process
 - a. Extension: The noun (or noun phrase) is placed in contexts where it no longer denotes an entity having third person reference but rather can be understood as standing for a concept of personal deixis, i.e. for a speech participant.
 - b. Desemanticization: The noun loses its lexical meaning in favor of a schematic, deictic function.
 - c. Decategorialization: The noun loses most of its nominal properties, such as the ability to inflect and to take modifiers.
 - d. Erosion: The noun may also lose in phonetic substance, i.e. become shorter.

The process sketched in (4) can be illustrated with a few examples from European languages involving honorific expressions (cf. (3a)). Spanish *Vuestra Merced* and Portuguese *Vossa Mercê*, both meaning ‘Your Grace’, were grammaticalized to polite second person pronouns (*Usted* and *Você*, respectively), Italian *La Vostra Signoria* ‘Your Lordship’ gave rise to the second person pronoun *Lei*⁷, or Dutch *Uwe Edelheid* ‘Your Nobility’, also developed into the second-person polite form *U* (see Head 1978: 185; Siewierska 2004: 224). In accordance with their origin as noun phrases, the resulting forms are likely to exhibit third-person agreement,

7. In Romanian, a similar development can be observed, leading to new address forms for all categories of person and number (Head 1978: 186).

even though they are now unambiguously markers of second-person deixis (see Comrie 1975).

All four parameters were involved in these developments. For example, the Portuguese possessive noun phrase *Vossa Mercê* underwent extension, desemanticization and decategorialization when it began to be used as a polite form of address for the king around 1460, but was soon extended to other social groups: Around 1490 its use was extended as an address form for dukes, in the 16th century it was further extended to the bourgeois population and in the 18th century it replaced the earlier *vós* as a polite and ceremonial form of address for singular referents, gradually undergoing erosion by being shortened to *você*.

An extremely advanced development can be observed in Brazilian and Angolan Portuguese, where *você* has largely eliminated the old second person singular pronoun *tu* as a general second person pronoun (Martelotta and Cezario in press; Merlan 2006: 222–224). In Brazil, the pronominalization of *você* began presumably at the end of the 18th century although it was only in the 20th century that it replaced *tu* almost everywhere⁸. Especially in informal speech, the form *você* is reduced to *ocê* and *cê*.

A similar process took place in Romanian, but unlike the cases just mentioned, there was hardly any erosion when the noun phrase *Domnia Ta* ‘Your Grace’ was grammaticalized to the second person address form *dumneata* (also: *Domnia Voastra* ‘Your (PL) Grace’ > *dumneavoastra*). This process was presumably associated with the rise of the Romanian feudal society between the 10th and the 14th centuries and the synthetic form of the pronoun is documented already in the 16th century (Merlan 2006: 222–226).

Which of the domains listed in (3) is recruited is much a matter of areal and cultural forces and it also determines to some extent which category of personal deixis may arise. For example, honorific expressions of the kind illustrated above (cf. (3a)) are most likely to be used for and to grammaticalize into second person reference. Generic nouns for human beings (3e), by contrast, are equally likely to give rise to first person plural pronouns.

Two examples may suffice to illustrate the latter pathway. In the Labwor language of northeastern Uganda there is an inclusive first person plural pronoun *jò* ‘we (including you)’ which, as both internal reconstruction and evidence from closely related Western Nilotic languages suggest, is a grammaticalized form of the noun *jò* ‘people’, e.g.

8. It is still possible to find the current use of the pronoun *tu* in certain specific interactional settings with third person singular agreement on the verb, as with *você* in Brazil.

- (5) Labwor (Western Nilotic, Nilo-Saharan; own field data)

én ʒnènɔ jɔ.
 3.SG see.PFV.3.SG people
 'He has seen us (all).'

Similar examples can be found in a range of other African languages, covering three of the four African language phyla: The Niger-Congo family is represented with languages such as Kono, Susu and the Cangin languages of Senegal, Nilo-Saharan with Ngiti, and Khoisan with !Xun (see Heine and Song in press for examples). But there are also examples from other parts of the world, including European languages. An example is provided by Portuguese,⁹ where the definite feminine singular noun phrase *a gente* 'the people' has been grammaticalized to a first person plural pronoun, to some extent also to a first person singular pronoun¹⁰—a process that is said to have started in the 17th century. In its lexical form, *gente* 'people' is a full-fledged noun, which can be inflected and modified; as a pronoun *a gente* is decategorizedized (with the preposed definite article *a* 'the'), e.g.

- (6) European Portuguese

A gente fala mais tarde.
 (the people talk more late)¹¹
 'We'll talk later.'
 (Merlan 2006: 230)

But generic nouns for human beings (3e) may also develop into third person pronouns; the following example may suffice to illustrate this pathway (for more examples see Heine and Song in press). The Central Sudanic language Lendu has suppletive forms for the noun 'man': The singular stem *ke* 'man' changes to *ndrú* or *kpa* 'people' in the plural and this suppletism has been retained in the grammaticalization from noun to personal pronoun: the third person singular pronoun is *ke* and the third person plural pronoun *ndru* or *kpa*:

- (7) Lendu (Central Sudanic, Nilo-Saharan)

ma- zhi ndrú. ke zhi kpa.
 1.SG-love 3.PL 3.SG love 3.PL
 'I love them.' 'He loves them.'
 (Tucker 1940: 392)

9. For a similar example involving French on, see Section 2.5 below.

10. In addition, a *gente* also shows uses of an indefinite pronoun corresponding to French *on* or German *man* (Merlan 2006: 233).

11. No glosses provided by the author; glosses are ours.

In the languages of Southeastern Asia, virtually all of the nominal domains listed in (3) were drawn on (see Cooke 1968 for a wealth of data; see also Heine and Song in press). For example,

- in accordance with (3c), nouns for ‘servant’ or ‘slave’ (e.g. Thai *khâa*, Burmese *tyunv* and Vietnamese *tôi*) serve to express first person reference and nouns for ‘master’ or ‘lord’ have been grammaticalized to second person reference
- in accordance with (3d), Thai, Burmese, or Vietnamese nouns for ‘teacher’, ‘master’, ‘doctor’ and ‘abbot or priest (of highest rank)’ have been extended metonymically to signal high status, most of all as second person address forms, but ‘teacher’ also for first person reference (Cooke 1965: 260).

In Khmer, “every common noun which descriptively denotes a status, occupation, or family relationship, is available as a pronoun of address but is by no means limited to second person deixis” (John Haiman, p.c.). Among the most common of these address forms are kinship terms and other status terms. Thus, in the following example, Khmer *knia*, which otherwise is a noun meaning ‘companion’ or a reciprocal pronoun, serves as a first person pronoun:

- (8) Khmer
knia *cih* *kaw:ng*.
 companion ride bike
 ‘I ride a bike.’
 (John Haiman, p.c.)

2.2. *Spatial deixis*

While clearly less common than nominal concepts, space nevertheless provides an important template for understanding and expressing personal deixis. This applies in particular to the spatial deixis of demonstratives, which is characterized by Greenberg thus:

Demonstratives [...] identify the referent by a local relation to the point of reference. These reference points are in many languages the persons themselves. The Serbs say *ovo meni, to tebi i ono njemu* ‘this for me, that for you and that yonder for him’, thus explicitly relating the three demonstratives to the three persons. (Greenberg 1986: xx)

Spatial deixis is treated in some works somehow as belonging to the same cognitive domain as social deixis, or as being inherently interconnected with social deixis in some way or other. While such works have contributed to a better understanding of the cognitive and social dimensions making up honorific systems, we will treat spatial and social dis-

tance as distinct domains. The former frequently serves as a model to conceptualize the latter, while the reverse, whereby spatial distance is expressed by means of categories of social distance, appears to be uncommon. Spatial deixis is used primarily to express the contrast between first and second person referents in terms of demonstrative or adverbial deixis, where the first person is encoded by means of proximal deixis ('here') and the second person by means of distal deixis ('(over) there'), as in the following example from Korean:

(9) Standard Korean

i jjog-eun gwaenchan- eunde geu jjog- eun
 this side-NOM good- CONN that side- NOM
eotteo- seyo?
 how- END
 'I am OK, and how about you?'
 (Song 2002: 14)

Spatial deixis has given rise to fully grammaticalized second person pronouns. The Japanese second person pronoun *anata*, for example, originates in a spatial deictic noun signaling roughly 'over there' (or 'that part') in Late Old Japanese, having been used as a third person marker ('person over there') in Early Modern Japanese, before it shifted to a second person pronoun around 1750 (Traugott and Dasher 2002: 230). This example is suggestive of a chaining process whereby expressions for spatial deixis give rise first to third person markers which again develop into second person pronouns (see 2.5).

Beyond the modality of spoken language it is most of all in sign languages that the deictic space has been recruited as a conceptual template for structuring personal deixis: Pointing gestures (and eye gaze) provide a convenient source for concepts of both spatial and personal deixis, and the pathways of grammaticalization proposed by Pfau and Steinbach (2006; see also Pfau and Steinbach in press) include the following:

(10) Sign languages (Pfau and Steinbach 2006)

pointing > locative > demonstrative > personal > agreement
 gesture pronoun pronoun marker

2.3. *Identifiers*

Under the term 'identifier' we subsume three kinds of pronominal concepts, namely reflexives, intensifiers¹², and identity pronouns ('the

12. The term "intensifier" is also associated with a number of different notions in linguistics. We are using the term here strictly in the sense defined by König and Siemund (2000), and described below.

same'). What the three have in common is that they all presuppose some entity whose referential identity has been established in previous discourse, and that they tend to undergo the same kind of grammaticalization process. In much of the existing literature, reflexive forms (e.g. *Paul killed himself*) and intensifiers (*John himself killed her*) are not distinguished terminologically, with the term 'reflexive' being used indiscriminately for both, especially since intensifier pronouns tend to be referred to as 'emphatic reflexives'. It therefore remains frequently unclear which of the two is involved in the creation of personal pronouns; it would seem, however, that in most such cases it is intensifiers, rather than reflexives, that are used.

An example of an identity pronoun giving rise to a second person address form is provided by the German identity pronoun *dieselben* ('the same ones'), cf. (11). The pronoun surfaced in the middle of the 18th century as an anaphoric marker of the highest level of honorification. But *dieselben* had a relatively short lifespan, declining in the 19th century, even if it was still mentioned occasionally around 1900. Furthermore, it was largely confined to plural addressees and to uses as an anaphoric marker referring to some antecedent title of dignity (Simon 1997: 274–275), and it was never conventionalized to a full-fledged second person pronoun (Simon 1997: 268; 2003: 127).

- (11) 18th century German
Ich bitte dieselben.
 I ask the.same.ones
 'I ask you.' (Maximally polite form of address)
 (Simon 2003)

Another example, involving contact-induced grammaticalization in Basque, is reported by Haase (1992). While its Romance neighbor languages Gascon and French have third person pronouns (Gascon *eth* 'he', *era* 'she'; French *il* 'he', *elle* 'she'), Basque has no dedicated third person pronouns. In order to develop an equivalent third person pronoun and thereby achieve translational equivalence with their Romance model languages, Basque speakers are developing their identity pronoun *ber-* 'same, -self' into a third person pronoun (Haase 1992: 135–137; see also De Rijk 2008: 114–115, 794–795).

In a number of languages, nouns for 'body' have given rise to markers of personal deixis. As we observed in Section 2.1, however, such cases do not lead straight from noun to personal pronoun but rather involve an intermediate stage of an intensifier or reflexive concept, hence there is the following chain of grammaticalization:

- (12) 'body' > intensifier and/or reflexive > personal pronoun.

Such a chain can be observed in some languages of Southeastern Asia. Thus, Thai *tuá* is not only a noun for ‘body’ but also means ‘self’ and serves as a second person form in specific contexts, and Vietnamese *minh* ‘body’ is not only used as a reflexive but also as a first and a second person pronoun (Cooke 1968: 17, 112).

There are a fairly large number of languages that are claimed to have developed an intensifier or reflexive marker into some kind of a marker of personal deixis.¹³ Siewierska (2004: 224–225) notes that in the languages which have other honorific forms, the reflexive is typically considered to be especially respectful and deferential and she mentions Kannada *taavu* (in the plural), Marathi *aapaṅ* and Maithili *apane*. But she also mentions the case of Malayalam, where the reflexive *taan-* ‘self’ is used with second-person reference for an addressee with whom one is on equal or familiar terms.

What the motivations underlying the identifier strategy are is a subject matter that requires much further research. Head (1978: 181) attributes the use of intensifiers (‘reflexives’ in his terminology) to their function of signaling emphasis: English *It’s John himself*, he argues, is more emphatic and thus may attach more importance to the referent than *It’s John*. Intensifiers can in fact contribute to this end. According to König and Siegmund (2000: 44–45), (adnominal) intensifiers “evoke alternatives to the referent(s) of the NP to which they are adjoined and characterize these alternatives (Y) as [in the] periphery or entourage of the referent(s)”. Conditions for the use of adnominal intensifiers include referents (X) that have a higher position than Y in a hierarchy, that are more significant than Y in a specific situation, or that are the center of perspective. Thus, in the English sentence *The chancellor himself was surprised at the results*, an alternative is inferred based on the fact that the chancellor (X) is high in rank compared to other possible referents.

It is this potential of intensifiers that appears to be exploited in their grammaticalization to second person pronouns: The speaker avoids addressing the hearer directly and portrays him as being more significant or central vis-à-vis alternative referents. Since intensifiers also provide the means of avoiding direct deictic reference to the hearer (see below), they constitute an almost ideal source concept for the grammaticalization of second person pronouns.¹⁴

13. In addition to the languages mentioned above, these languages are Bengali, Hindi, Japanese, Kannada, Kashmiri, Khmer, Korean, Maithili, Malayalam, Marathi, Tetelcingo Nahuatl, Imbabura Quechua and Turkish.

14. The question then remains what accounts for the ability of identity markers to serve as conceptual sources for second person pronouns. The answer presumably has to do with

The data that are available to us suggest that avoidance plays an important role in at least some of the languages that have drawn on this strategy. Shibatani (1985: 837) suggests that the connection between reflexives and honorifics lies in agent defocusing: “Honorific speech is characterized by the avoidance of the singling out of an agent, be it the speaker, addressee or third party ... The marking of the absence of an agent in the spontaneous construction may therefore be adopted for the agent defocusing in honorification” (quoted from Siewierska 2004: 227).

Siewierska (2004: 227) doubts, however, whether this holds for the honorific use of pronominal reflexives, e.g. in the Dravidian and Indo-Aryan languages. She draws attention to a possible connection between spontaneous constructions, pronominal reflexives and honorification via the use of intensifiers in their ‘alone’ or ‘without any assistance’ readings: “If used with inanimate subjects, such reflexive pronouns may produce a reading whereby the event has come about effortlessly or spontaneously [...] This spontaneous reading of the pronominal reflexive would thus allow for the extension to the agent defocusing or impersonalizing interpretation, which underlies much honorific usage” (Siewierska 2004: 227).

When grammaticalized to a second person pronoun, the intensifier, reflexive or identity pronoun is likely to develop into an invariable form that may no longer refer to categories other than second person deixis. The reflexive form *ap* of Hindi and Urdu, for example, occurs in all categories of person and number; as an address form; however, where it expresses the most respectful pronominal form of address, it has only second person reference (Head 1978: 180). With an increasing degree of grammaticalization, identifiers may also undergo other function-specific changes. In general, reflexives used as person markers show third person agreement; however, once they turn into markers of personal deixis they may adapt their agreement behavior to their new grammatical function.

2.4. *Pluralization*

Grammaticalization normally involves a process leading from one kind of category to another. But it may as well involve one and the same general category, where one member of a paradigm is conceptualized and expressed in terms of another member. Within the category of personal pronouns, two of the strategies mentioned above are of the latter kind; one of

the following observations. Identity markers and intensifiers have to some extent overlapping functions, as has been pointed out by a number of authors, and in a number of languages they are etymologically the same. For example, in German the root *selb-* is found both in the intensifier *selbst* and the identity adjective *der-selbe* ‘the same (masculine)’.

them, pluralization,¹⁵ is looked at in the present section, and the second one, leading from third person to second person pronouns, in Section 2.5.¹⁶

In Section 2.2, we described pluralization, whereby second or third person plural pronouns are used for second person singular address, as one of the most salient strategies in the rise of new personal pronouns. Pluralization has been portrayed as being of universal significance, to be observed in all major parts of the world, even if it is rarely found in the languages of the Americas (Head 1978: 158; Siewierska 2004: 216–221).

Paradigm examples of pluralization are encountered in highly stratified societies. The last emperor of Ethiopia, Haile Sellasie, spoke of himself as *əŋŋa* ‘we’ (Zealelem Leyew; p.c.). A similar situation obtained in other highly stratified societies, such as Cambodia, Thailand and the Roman Empire. Thus, in Khmer, the first person plural pronoun *jeu:ng* ‘we’ also serves for the royal first person singular ‘I’ (John Haiman, p.c.), and a similar case has been reported in Thailand: In Thai, the first person plural pronoun, *raw* (‘we, us, our’), being the plural of most first person singular forms that exist in the language, serves as a singular form for the king speaking to subjects in public address, and it is also used more generally by a superior to an inferior, especially an employer talking to an employee, even if the last usage does not appear to be entirely conventionalized, considered by some to be a little distant (Cooke 1965: 25).

2.4.1. *Pluralization in European history.* It may be worth looking at the case of the Roman Empire in greater detail since it illustrates the effects that the pluralization strategy had for a number of European languages. Most western and central European languages have a grammatical distinction between two forms of second person singular reference. Following the tradition established by Brown and Gilman¹⁷ (1968; for more details, see there) we refer to the two forms respectively as the *T*-form (cf. Latin *tu* ‘you.SG’) and the *V*-form (cf. Latin *vos* ‘you.PL’);

15. Note that we use ‘pluralization’ here in a specific sense, ignoring all the other usages of this term. Note further that with this term we are referring not only specifically to plural forms but also include dual and other non-singular categories.

16. We are not only restricted specifically to plural forms but also include dual and other non-singular categories. Siewierska (2004: 220) observes that in languages that distinguish between dual and plural, it is typically not the dual but the plural that is used for honorific singular address, even if she found some notable exceptions in Oceanic languages.

17. Note that the theoretical assumptions made by these authors have not found general approval in subsequent sociolinguistic research (see e.g. Braun 1988). This however is an issue that need not concern us here.

French *tu* vs. *vous* is an example of this distinction. The rise and subsequent development of this distinction covers a history of nearly two millennia; in the following we are restricted to a few salient traits of it.

In the Latin of antiquity there was only the second person singular pronoun *tu*. The Roman emperor sometimes spoke of himself as *nos* 'we', and he was addressed by means of the plural pronoun *vos*, which thereby acquired a new, additional meaning, namely denoting a singular referent.¹⁸ The result was the rise of a reverential *vos*, much later imitated by some other European societies. The medieval German emperors replicated this principle in their official documents written in Latin and in the 9th century, German *ir* (you.PL) began to be used in the same way (Lockwood 1968: 61). The first context extension involved the spread of the Latin plural form *vos* from the emperor to the nobility. This use was adopted and further extended in post-Roman European societies: In medieval Europe, generally, the nobility said *T* to the common people and received *V*.

Context extension reached in fact a new dimension in medieval Europe: The use of the *V*-form spread from higher social status more generally to people having control over others, and from principal statal power to the family: "[...] the master of a household said *T* to his slave, his servant, his squire, and received *V*. Within the family, of whatever social level, parents gave *T* to children and were given *V*" (Brown and Gilman 1968: 255). In addition to the non-reciprocal pronominal usage there was also a reciprocal one: In medieval European societies, pronominal address was reciprocal between equals, where equals of the upper classes exchanged the mutual *V* and equals of the lower classes exchanged *T*.

Subsequently, there was a transition from vertical distance to horizontal distance, whereby social hierarchy was replaced by psychological, affective components of proximity and familiarity (see Held 1999). The *V*-form loses its function as a marker of asymmetric social relations and is generalized in specific domains as a marker of symmetric and reciprocal relationship, that is, speaker and hearer invariably exchange either *V* or *T*. Parameters for the choice between *V* and *T* are now social or kinship relationship, degree of solidarity or familiarity. This stage is characteristic of language use, e.g. in modern France.

A final stage in the development of context extension was reached when the *V*-form was generalized, being extended to all contexts where

18. That it was self-reference that introduced the *V* form is suggested by statements such as the following: 'A man who spoke of himself as *nos* was accordingly addressed as *vos*' (Lockwood 1968: 61).

the *T*-form was used—with the effect that the latter was given up in favor of the *V*-form. English has reached this stage, where the *V*-form *you* has entirely ousted the earlier *T*-form *thou*. Whether a process like the one that happened in English takes place or, alternatively, whether there is a process in the opposite direction, leading to the extension of the *T*-form at the expense of the *V*-form, as appears to be happening in some European languages, depends on the particular socio-political forces characterizing the situation in which the *T/V*-distinction is made (Fritz Newmeyer, p.c.).

Pluralization can be found in a wide range of languages and in most cases it serves the expression of respect and/or social distance. As the discussion by Siewierska (2004: 216–221) reveals, however, its functions are complex, that is, the extension from plural to singular referents can follow special pathways. Among the Bemba, a Bantu-speaking society of Zambia, for example, adults are addressed by means of second or third person plural forms, but the second person singular pronoun is used with paramount chiefs (Gregersen 1974: 53).

That pluralization is a complex strategy can be seen e.g. in the fact that in first-person reference there is not only a plural of majesty but also a plural of modesty. In classic 18th century novels in Chinese, individuals of inferior status such as women, servants, or children, give the first person plural for self-reference to persons of superior status (Lee 1999); one may also wish to mention the ‘editorial we’ in European languages, at least in one of its functions. Note also that in Sierra Popoluca of Mexico, the first person plural inclusive is claimed to be the normal form of polite self-reference while the first person singular is considered as authoritarian (see Siewierska 2004: 218–219 for more details and insightful discussion).

A number of explanations have been volunteered to account for pluralization. According to one, it is interpreted as a natural metaphor for social power, e.g. in terms of a conceptual metaphor MORE IS BETTER (Lakoff and Johnston 1980; Siewierska 2004: 217); according to another explanation, it is argued to be a means of redressing the negative face wants of the addressee, a plural form of address being less face threatening than a corresponding singular form (see Brown and Levinson 1987: 198 for arguments).

While the conceptual transfer pattern from forms for plural (or dual) referents to singular referents is generally unidirectional, Siewierska (2004: 221) found a counterexample: For 18th century Chinese, a case has been reported where speakers of high social status discuss referents of inferior social position, whom they consider as unimportant, using singular forms for non-singular reference. Note however that 18th century Chinese is also unusual in some other ways (see above); for example, plu-

ralization as a denigrating device applies not only to first person but also to second and third persons; for example, the third-person plural form *ta-men* instead of the third person singular form *ta* indicates that the referent is negligible or of little importance (Siewierska 2004: 219, based on Lee 1999).

2.4.2. *A case study of pluralization: Sango.* The pluralization processes that we discussed above are old, dating back to past centuries. But such a process can happen any time, and it does not presuppose specific social situations. The process looked at in the following example occurred fairly recently, within the last fifty years, and it arose in the context of a largely egalitarian modern African society. Sango, a Ubangian language of the Niger-Congo family, is the national language of the Central African Republic, and it has also been declared an official language side by side with French. Arising as a pidgin at the end of the 19th century, it acquired a substantial number of mother tongue speakers in the course of the last decades and in roughly the same period the language underwent a number of structural changes. One of these changes involved the rise of a new functional category via the extension of the second person plural pronoun *álà*, corresponding to the European *V*-form, to a deferential second person singular pronoun (Samarin 2002). What makes this a particularly interesting case is that it allows us to study such a change in its *status nascenti*. The rise of the category occurred after the Central African Republic attained its independence; there is no evidence for its existence prior to the 1960s.

Perhaps the main characteristic of this process can be seen in the fact that it seemingly does not correspond significantly to any fixed sociolinguistic principles: Essentially any Sango speaker can give *álà* to anyone else instead of the traditional second person singular pronoun, the *T*-form *mò*. The only clear constraint in the use of the strategy concerns children below 13 years of age, who do not address younger siblings with *álà* (Samarin 2002: 307). Nevertheless, there are degrees of probability in the use of the strategy, based on the demographic variables sex and age: Female addressees are generally more likely to receive *álà* than males, and the probability that grandparents and parents receive *álà* is the highest, being in the range between 77 and 89 per cent of Samarin's data.¹⁹ The social categories least affected by the process are younger siblings and friends, with whom *álà* is used with less than 25 per cent probability.

19. Unfortunately, Samarin (2002) does not make it clear what the corpus is on which these figures are based.

To conclude, Sango speakers have acquired a new functional category for which previously there was no equivalent in the language. Being a young category, it has not yet been fully conventionalized: Its use remains largely optional, being determined to some extent by social and demographic variables. It is available as a rhetorical option that may be used or ignored in accordance with the speaker's preferences and communicative intentions.

2.5. *Shift in deixis*

Another pathway of grammatical development concerns the shift from one category of personal deixis to another. There are two major manifestations of this general pathway: On the one hand, it involves a shift from definite noun phrase to second person pronoun; on the other hand, it leads from generic human noun ('people' or 'person(s)') or an impersonal pronoun to first person plural pronoun.

2.5.1. *The creation of new second person pronouns.* One example was mentioned in Section 2.2 in connection with spatial deixis: The Japanese second person pronoun *anata* 'you' was a noun for spatial deixis in Late Old Japanese and turned into a third person marker ('person over there') in Early Modern Japanese before it underwent a second shift from third to second person pronoun around 1750.

Another example concerns developments in German that were already alluded to in Section 2.1 with reference to the *noun-ellipsis channel*. This is a more complex process, whereby a nominal argument is no longer overtly mentioned in a clause but there remains a third person anaphor that assumes the function of a personal pronoun. During the transition from the 16th to the 17th century, nominal forms such as *der Herr* ('the gentleman') *die Jungfer* ('the damsel'), or *der Vater* ('the father') abounded in German for second person reference and underwent semantic extension, to the point that it was frequently unclear whether these expressions had second person or third person reference. Not uncommonly, the earlier polite plural form *ir* 'you (PL)' or its possessive equivalent *ewer* was used as an anaphoric pronoun, as in the following example:

- (13) 17th century German (Herzog Heinrich Julius von Braunschweig, 1594)

Der Juncker hat ja nach mir geschickt, Was ist
 the squire has yes after me sent what is
ewer beger?

2.PL.POSS desire

'You (the squire) have sent for me; what do you want?'

(Simon 2003: 109)

But more generally, nominal subject referents were taken up anaphorically by the corresponding personal pronouns *er* (3.SG.M) and *sie* (3.SG.F), respectively, agreeing with their nominal antecedent in gender. Subsequently there was a gradual transition from agreement based on grammatical gender to agreement based on semantic gender, and from nominal arguments with anaphoric coreference to pronominal arguments not requiring any nominal antecedent. For example, the noun *Fräulein* ('miss, young lady') has neuter gender and hence took the neuter pronoun *es* as an agreement marker. After roughly 1670, however, it was replaced by the feminine pronoun *sie* ('she') when referring to a female participant, irrespective of the grammatical gender of the noun concerned, and it no longer required a coreferential nominal argument. Note that this development can be interpreted as an instance of *degrammaticalization*, in that a personal pronoun serving as a marker for grammatical agreement and requiring a nominal antecedent, gave way to a semantically defined independent personal pronoun not requiring a preceding co-referential noun (phrase); more data are required for a comprehensive analysis of this case.

The result thus was that with the ellipsis of the nominal forms, German acquired the personal pronouns *er* (3.SG.M) and *sie* (3.SG.F) as new second person politeness pronouns. The two new pronouns had a short life-span, however: around the middle of the 19th century they were lost²⁰ (see below).

A second development from third to second person took place slightly later in 17th century German, when a new strategy came to be employed for expressing 2.SG reference: Rather than generic classifying nouns or kin terms that were used earlier, it was now abstract honorific expressions denoting human traits or virtues that were recruited (cf. (3a) above), such as *Gnaden* ('grace'), *Ehren* ('honor'), or *Heilichkeit* ('holiness'), typically relating to salient status properties of the addressee. Thus, a teacher or scholar would be addressed e.g. as *Eure Weisheit* ('your wisdom'), or a prince as *Eure Majestät* ('your majesty'):

By appealing to the addressee's graciousness, mercy, majesty, holiness or other such flattering characteristics, the speaker brings about a similar sort of indirectness as with second person plural address: the addressee is referenced indirectly via associated entities. (Listen 1999: 57; quoted from Simon 2003: 111)

The forms were quickly grammaticalized, being commonly abbreviated in writings (e.g. *E.G.* for *Euer Gnaden* 'your grace') and extended to all

20. They survived, however, in their old function as third person pronouns, being part of the paradigm of personal pronouns of Modern High German.

argument slots, and their use became virtually mandatory in some speech situations (Listen 1999: 61; Simon 2003: 111). And as in the case of the earlier *er/sie*, the personal pronouns used to refer to nominal antecedents gradually turned from anaphoric to semantically defined pronouns. German abstract nouns have feminine gender and hence would be referred to anaphorically with the feminine pronoun *sie* ‘she’ but in the course of the 17th century the masculine pronoun became more common when the referent was male.

While the abstract nouns that assumed the function of 2.SG address forms had either singular or plural number, it was the 3.PL pronoun (*Sie*) that was gradually generalized for all nominal expressions²¹, and roughly from the beginning of the 18th century onward, *Sie* had won out against the singular pronouns and could occur without any nominal antecedent (Simon 2003: 114). Around 1800, the use of *Sie* was extended from the upper classes to the middle class and its generalization across different social strata could have been responsible for the gradual decay and disappearance of both *ir* and *er/sie* around the middle of the 19th century (Simon 2003: 121).

Another example of an evolution from third to second person concerns the Japanese second person pronoun *anata* ‘you’; we have looked at this case already in Section 2.2: Having its origin in a spatial deictic noun in Late Old Japanese, *anata* turned first into a third person marker (‘person over there’) in Early Modern Japanese before it underwent a second shift into a second person pronoun around 1750 (Traugott and Dasher 2002: 230), but we have no information on what motivated this shift.

2.5.2. New first person plural pronouns. The second pathway of deictic shift leads from an impersonal pronoun, having third person reference, to a first person plural pronoun. The following example concerns the development from the French impersonal pronoun *on* ‘one’ (itself historically derived from the Latin noun *homo* ‘person’, ‘man’) to the first person plural pronoun *on*.²²

While there have been attempts to trace the origins of the new personal pronoun *on* back to Middle French, or even to Old French, it was not in general use in the Paris area before the 19th century (Coveney 2000: 450), that is, it appears to be no more than two centuries old. In its modern

21. The reasons for this development are complex and are not entirely clear; see Simon (2003: 113–114) for discussion.

22. To simplify matters, we will focus on the pronominal clitics *nous* and *on* and ignore verbal personal inflections (i.e. what are technically known, respectively, as ‘4th person verbs’ and ‘3rd person verbs’).

function as a first person plural pronoun, *on* differs in a number of ways from other personal pronouns, including the older first person plural pronoun *nous*. The linguistic context characterizing this replacement is the use as a subject clitic pronoun, while the sociolinguistic context is described as working class French in particular and informal spoken French in general. In Coveney's (2000) text corpus of a quarter million words of informal metropolitan spoken French, *nous* is present only very marginally, the clearly preferred first person subject pronoun being *on*. Thus, the new pronoun has undergone massive extension. But the extension of the new pronoun is far from complete, as the old pronoun *nous* has retained its status in the following contexts:

- (a) when used as an object pronoun,
- (b) as a subject pronoun in non-clitic ('disjunctive') uses, where it can be stressed for contrastive or emphatic purposes,
- (c) in dislocated positions,
- (d) in cleft sentences where the matrix clause has a first person plural argument (e.g. *c'est nous qui...* 'it is we who...'), or
- (e) if followed by *-mêmes* '-selves' (i.e. *nous-mêmes* 'we ourselves'; Coveney 2000: 456).

In sociolinguistic terms, *nous* has survived in written as well as in spoken formal speech; it tends to be used as a 'prestige variant' for talking on formal topics (Blanche-Benveniste 1997: 23) or, more generally, when a speaker shifts from an informal to a formal style (Coveney 2000: 477). It has been claimed that *on* (a) implies some degree of familiarity where *nous* does not, that (b) it tends to be used for inclusive reference and *nous* for exclusive reference, or that (c) it is used to refer to a group 'seen from within' whereas *nous* refers to a group 'seen from the outside' (Blanche-Benveniste 1985: 208). While such observations are not without foundation, none of them is entirely convincing. After a detailed text analysis, Coveney (2000: 470, 464–465) comes to the conclusion that the choice between *on* and *nous* is not motivated fundamentally by semantic considerations. But perhaps more importantly, there appear to be quite a number of speakers who have entirely replaced *nous* by *on* as a subject clitic; hence, for such speakers no semantic distinction exists.

To conclude, this French example suggests that the emergence of a new personal pronoun is not necessarily motivated by language-internal requirements, that is, linguistic factors such as a gap in the system or any other structural 'pressure', nor by social or sociocultural requirements—we are not aware of any convincing evidence to the effect that French speakers really 'needed' a new pronoun. At the same time, we were not able to proffer any alternative account—functional or non-functional—

that would be of help for understanding the process; more research is required on this issue.

2.6. *Combining*

Finally, mention should be made of another kind of strategy for forming new categories of personal deixis, namely via the combining of already existing categories. Presumably the most common form of combining is by adding a number marker to the singular form of personal pronouns to form plural or other non-singular pronouns.

One language where this has happened to some extent is Vietnamese. By preposing the pluralizer *chúng* ‘they’ to some other pronouns, additional deictic meanings are created. Thus, Cooke (1968: 115) observes:

All personal pronouns proper except *minh* and *ta* are basically singular when standing alone, and are pluralized by the addition of preposed *chúng* [...] First person plural forms, all of which are personal pronouns proper, may either include or exclude the addressee. Thus, *chúng mình*, *chúng ta*, and also *mình* and *ta* in plural usage, include the addressee with the speaker and convey the meaning, I and thou, or I and you (plural); but other plural forms, namely *chúng tôi*, *chúng nó*, and *chúng tao*, are exclusive of the addressee and therefore denote I and he, I and she’, or ‘I and they’.

In a more consistent way than Vietnamese, speakers of Tok Pisin have designed their system of personal pronouns by forming non-singular pronouns by means of number markers: They grammaticalized the numeral *tu* ‘two’ to a dual marker, the numeral *tri* ‘three’ to a trial marker (apparently not widely distinguished), the marker *-fala* or *-pela* (<English *fel-low*) to a non-singular marker, and the combination *yu* ‘you’ plus *mi* ‘I’ to a first person inclusive marker; cf. Table 1.

Table 1. *Tok Pisin personal pronouns* (Jenkins 2002: 216).

Number and person	Singular	Dual	Trial	Plural
1 inclusive		yumi(tupela)	yumitripela	yumipela
1 exclusive	mi	mitupela	mitripela	mipela
2	yu	yutupela	yutripela	yupela
3	em	ol(tupela)	ol(tripela)	ol

While Vietnamese and Tok Pisin represent a common pattern of combining, there is one language that has drawn on a typologically unusual strategy of combining, as suggested by Brown and Dryer’s (2008) description. In the Walman language of the Torricelli family, spoken on the

north coast of Papua New Guinea, noun-phrase-conjoining conjunctions meaning ‘and’ have the appearance of transitive verbs, showing prefixal subject agreement with the first conjunct and suffixal agreement with the second conjunct. Note that this applies not merely to one form but to two, *-aro-* and *-a-*, which are generally interchangeable and semantically essentially the same.²³ Example (14) illustrates the relevant coordination construction with the ‘and’-verb *-a-*; the construction behaves syntactically like a simple noun phrase.

- (14) Walman (Torricelli)
Ako ru w- a- n muen mkie
 then 3.SG.F 3.SG.F and- 3.SG.M brother banana
y- oko- ø [...].
 3.PL take- 3.SG.F
 ‘Then she and her little brother took the bananas [...].’
 (Brown and Dryer 2008: 531)

Both the nominal subject and the nominal object can be omitted, in which case there are only verbal affixes representing subject and object or, respectively, the two conjuncts:

- (15) Walman (Torricelli)
Kurue [n- aro- n] y- r-
 but 3.SG.M- and- 3.SG.M 3.PL RECP/REFL
apar nyemi mlin.
 be.related.to friend true
 ‘But they ([he and him]) became true friends.’
 (Brown and Dryer 2008: 536)

Conjoining thus has given rise to what Brown and Dryer (2008: 537, 544) call ‘quasi-pronominals’, and the authors suspect that these items are grammaticalizing into pronouns:

These uses of the *and*-verbs with inflectional arguments are often most naturally translated into other languages using personal pronouns; for convenience we refer to them as QUASI-PRONOMINAL [...]. The quasi-pronominal uses of *and*-verbs, however, have the advantage of specifying more information; while *ri* is simply third plural, *naron* specifically denotes two males, *wan* specifically denotes one male and one female, and *nay* specifically denotes a male plus a plural set, [...]. They in fact provide a richer pronominal system than any true pronominal system we are aware of. (Brown and Dryer 2008: 537)

23. There are minor differences; for example, *-aro-* occurs only with third-person objects while *a-* also occurs with first- and second-person object prefixes.

3. The genesis of a system

All data that we have been looking at in the preceding sections concern the evolution of individual personal pronouns or groups of pronouns. But there is at least one example that shows how an entirely new system of personal pronouns can arise. This example concerns Calunga, commonly classified in Brazil as *falar africano*. Spoken mainly in the city of Patrocínio in the *Triângulo Mineiro* of the state of Minas Gerais, Calunga is used most of all by older Afro-Brazilian men and (less frequently) by African women, but there are also some white speakers of it (Byrd 2006). The language appears to be used exclusively as a second language; the primary language of all Calunga speakers is *caipira* (Brazilian Portuguese Vernacular). According to the description by Byrd (2006: 73), Calunga arose among slaves who first acquired Portuguese as their new language but later developed Calunga as a 'secret language'.

Calunga has a substantial lexicon that can be traced back to Bantu languages of southwestern Africa, most of all to the Angolan language Kimbundu, considerably less to Umbundu and Kikongo. Since all three languages are spoken in Angola (though Kikongo is more widespread in the Democratic Republic of Congo), there is reason to assume that the speakers who were crucially involved in the development of Calunga overwhelmingly originate from Angola. But in addition to this African vocabulary, there is also a substantial Portuguese lexicon.

While linguistic innovations in Calunga are to be found most of all in the lexicon, there are also new functional structures. The most noteworthy structure concerns personal deixis: Calunga speakers created a new system of personal pronouns by grammaticalizing lexical items to markers of personal deixis. As Table 2 shows, this system has little in common with any of the languages that contributed to the growth of this language. Nevertheless, the earlier pronouns, matching those of *caipira*, are still available as optional variants, in particular the plural pronouns *nóis* 'we, us' and *(o)cês* 'you (PL)', which can be used freely instead of the 'genuine' Calunga pronouns listed in Table 2.

The system is in some ways unique; to our knowledge, there is no other pronoun system exhibiting a similar range of conceptual distinctions (cf. e.g. Forchheimer 1953; Ingram 1978; Mühlhäusler and Harré 1990; Haase 1994; Cysouw 2003; Siewierska 2004). The main characteristics of the pronoun system summarized in Table 2 can be described thus: First, it contrasts with that of most other languages in that the largest number of conceptual distinctions is found in the first person singular, where four different referential concepts are distinguished. Note that crosslinguisti-

Table 2. *Calunga personal pronouns* (optional elements are omitted; B = black referent; F = feminine; M = masculine; PL = plural; Byrd 2006: 114).

	Singular form	Literal meaning	Plural form	Literal meaning
1 (default)	camano-cá	person-here	os camano-cá	the.PL person-here
1.B	camano-fú	person-black		
1.B.M	umbundo-cá	black.person-here		
1.F	ocai	woman	as/os ocai	the.PL woman
2 (default)	camano-aí	person-there		
2.M			os camano-aí	the.PL person-there
2.F	ocai	woman	as/os ocai	the.PL woman
3.M	camano		os camano	the.PL person
3.F	ocai		as/os ocai	the.PL woman

cally first person singular tends to be the least differentiated category; for example, in their survey of 110 languages, Haberl and Hanson (n.d.: 2) did not find a single language having gender marking on first person singular pronouns.

Second, the number of deictic distinctions is reduced with female referents: There is only one feminine pronoun for second and third person. Third, it shows a gender distinction masculine vs. feminine in all three persons, while none of the possible ‘substrate’ languages, i.e. the Angolan Bantu languages Kimbundu, Umbundu, and Kikongo, has any grammaticalized sex-based distinction. And fourth, it contains a category that is hard to find in other languages, namely what Byrd (2006: 113) calls “ethnicity”: The forms *camano-fú* and *umbundo-cá* (or *imbundu cá*) refer to black referents and are used if the speaker “wishes to emphasize his ethnicity” (Byrd 2006: 113). That these pronouns contrast with those of any Portuguese variety spoken in Brazil can be shown with the following example, where the pronoun *umbundo-cá* (in its variant form *imbundu cá*) appears as a possessive attribute, encoded like a nominal possessee:

- (16) Calunga *fii* *du* *imbundu* *cá* ‘my children’
 child of black.person here
- Portuguese *meus* *filhos*
 my.PL children

The flexibility of Calunga discourse structure can be illustrated with the following example, which shows e.g. that deictic reference need not be specified as long as the context provides sufficient clues on reference identity:

(17) Calunga

O injó du camanu é aonde?
 the house of person is where
 ‘Your house is where?’

O injó du camanu é [...] por lá.
 the house of person is over there
 ‘My house is [...] over there.’
 (Byrd 2006: 189)

To conclude, the Calunga system of personal pronouns is a new one, one that emerged in the course of the last centuries and whose structure cannot be related to any of the languages that the Calunga speakers were familiar with. While being in a number of ways typologically peculiar, the system nevertheless is in accordance with the strategies that we outlined in Section 2: First, the system is the result of canonical processes of grammaticalization, whereby lexical material was recruited for the expression of grammatical functions of personal deixis. And second, grammaticalization involved two of the main strategies that we found to have been at work in the rise of other forms of personal deixis: It relied on the one hand on the noun strategy (2.1) whereby generic nouns such as *camano* ‘man’ and *ocai* ‘woman’ were drawn on; on the other hand it also made use of spatial deixis (2.2) by transferring the locative adverbs *ai* ‘there’ and *cá* ‘here’ from the domain of physical space to that of social space.

In accordance with what grammaticalization theory would predict, the Calunga case can be taken to be representative of what people do when they design new constructions: They create the new grammatical patterns by extending existing constructions to new contexts (cf. the extension parameter of (1)), and in doing so, they draw on a small pool of conceptual templates. In the example that we were concerned with in this section, these templates consisted of and were restricted to two of the conceptual sources listed in (2): human nouns (2a) and basic concepts of spatial orientation (2b).

4. Conclusions

The general motivation underlying the pathways described in this paper can be seen in human activity aimed at finding optimal ways of saying what is both socially appropriate and most advantageous for the speaker in a given sociolinguistic context. This motivation induces speakers to present the participants of a communicative act in a specific way. In some situations it may be desirable to avoid addressing the speaker directly; consequently, speakers use rhetorical strategies that are construed

as instrumental in responding to such situations, drawing on honorific noun phrases, identifiers, or distinctions of spatial deixis in search of an expression for second person reference that is felt to be most appropriate in a given social situation. The level of creativity that speakers exhibit when designing new forms of personal pronouns is apparent especially in Section 3, where we discussed a new system of personal pronouns that former African slaves have created in their new homeland of Brazil. This creativity consists essentially of using general principles of conceptualization to combine existing linguistic material in novel ways.

Findings made in this paper suggest, first, that personal pronouns are not primitives of grammatical evolution; rather, there is diachronic evidence to show that they often derive from other conceptual domains. As we saw in Section 2, there is a limited pool of conceptual sources that speakers tend to recruit in order to create new forms of personal pronouns. These sources and the main categories of personal pronouns derived from them are listed in Table 3. As Table 3 suggests, second person pronouns have the largest number of sources, while for first person plural pronouns we found essentially only one source, namely indefinite pronouns (like French *on*) and nouns meaning ‘person(s)’ or ‘people’, where the former are historically derived from the latter.

Table 3. *The main conceptual sources of personal pronouns.*

Source	Third person	Second person	First person	
			Singular	Plural
Nominal concepts	+	+	+	+
Spatial deixis	(+)	+	+	
Identifiers	+	+		
Pluralization		+		
Shift in deixis		+		

And second, these findings also suggest that the evolution of personal pronouns is essentially unidirectional. For example, personal pronouns commonly derive from nominal expressions, deictic demonstratives, etc. Conversely, we are not aware of any examples where personal pronouns turned into lexical nouns or demonstratives.²⁴ And there are examples of

24. Note, however, that personal pronouns can give rise to new demonstratives *when combined with some other element*, such as a demonstrative stem. But such developments do not affect the unidirectionality principle.

shifts from third to second person reference (see 2.5.1), but we were not able to identify any case where this process went in the opposite direction. While we do not wish to rule out the possibility that there are counterexamples to the unidirectionality principle, we maintain that they would be exceptional.²⁵

But in concluding we must draw attention to a problem that we did not confront in this paper. We were using grammaticalization theory as a tool to describe the evolution of personal pronouns and, in fact, most of the phenomena characterizing this evolution are in accordance with and can be described appropriately with reference to parameters of grammaticalization. Most importantly, as we just noticed, the diachronic trajectories yielding personal pronouns appear to conform to the unidirectionality principle.

Nevertheless, there are also remarkable differences compared to canonical instances of grammaticalization. First, grammaticalization is commonly described as a process leading from concrete, referential meanings to more abstract, relational meanings. While personal pronouns can be said to be relational in that they are defined with reference to the deictic situation in which they are used, their meaning is not abstract—at least not in the sense used in grammaticalization theory. Second, we encountered a case of degrammaticalization in Section 2.5.1, in that in 18th century German, third person singular pronouns serving as markers of grammatical agreement and requiring a nominal antecedent, gave way to a semantically defined independent personal pronoun not requiring a coreferential nominal argument.

But there is an even more serious challenge to grammaticalization theory when deictic shift is concerned (see 2.5). Grammaticalization can be understood as a process that entails a transition from one morphological or syntactic paradigm to another. Such a shift can in fact be observed in the case of three of the sources that we observed in Sections 2 and 3, namely those of nominal concepts, spatial deixis and identifiers. But, obviously, there is no shift in paradigm when either pluralization (2.4) or shift in deixis (2.5) is involved: What happens rather is that a member of the paradigm of personal pronouns changes into another member of the same paradigm. Thus, with the extension of the erstwhile English plural pronoun *you* to also serve as a singular pronoun, *you* continued to be part of the paradigm of English personal pronouns. Note further that deictic shift involves the parameter of extension but not necessarily any of

25. One possible counterexample, referred to as ‘degrammaticalization’, was pointed out in Section 2.5.1; see below.

the other three parameters of grammaticalization that we proposed in (1) of the introductory section.

We are thus left with the question of whether some, or all, of the evolutionary trajectories that led to personal pronouns fall beyond the scope of grammaticalization theory. Alternatively, perhaps the evolution of personal pronouns requires an expanded understanding of grammaticalization processes. More research is required on this issue.

Abbreviations

ABS = absolute pronoun; ACC = accusative; B = black referent; CL = classifier; CONN = connective; DU = dual; EX = exclusive; END = ending; F = feminine; H = honorific; HAB = habitual; HUM = human; IN = inclusive; INTR = interrogative particle; M = masculine; N = neuter; NOM = nominative; O = object; OBJ = object; PL = plural; POSS = possessive; PRES = present; RECP = reciprocal; REFL = reflexive; SG = singular; SUB = subject; TOP = topic; TR = trial; 1, 2, 3 = first, second, third person.

References

- Blake, F. R. 1934. The origin of pronouns of the first and second persons. *The American Journal of Philology* 55. 244–248.
- Blanche-Benveniste, C. 1985. Coexistence de deux usages de la syntaxe du français parlé. *Actes du XVIIe Congrès International de Linguistique et Philologie Romanes, 1983*, 7. 203–214.
- Blanche-Benveniste, C. 1997. *Approches de la langue parlée*. Paris: Ophrys.
- Braun, F. 1988. *Terms of address: Problems of patterns and usage in various languages and cultures*. (Contributions to the Sociology of Language 50). Berlin.
- Brown, L. & M. S. Dryer. 2008. The verbs for ‘and’ in Walman, a Torricelli language of Papua New Guinea. *Language* 84(3). 528–565.
- Brown, P. & S. Levinson. 1987. *Politeness: Some universals in language usage*. Cambridge: Cambridge University Press.
- Brown, R. & A. Gilman. 1968. The pronouns of power and solidarity. In J. A. Fishman. *Readings in the sociology of language*, 252–281. Berlin: Mouton de Gruyter.
- Bybee, J. L., R. D. Perkins & W. Pagliuca. 1994. *The evolution of grammar: Tense, aspect and modality in the languages of the world*. Chicago: University of Chicago Press.
- Byrd, S. E. 2006. *Calunga, an Afro-Brazilian speech of the Triangulo Mineiro*. Ann Arbor, MI: UMI Services.
- Campbell, L. & R. Janda. 2001. Introduction: Conceptions of grammaticalization and their problems. *Language Sciences* 23(2/3). 93–112.
- Comrie, B. 1975. Polite plurals and predicate agreement. *Language* 51. 406–441.
- Cooke, J. R. 1965. *Pronominal reference in Thai, Burmese, and Vietnamese*. Ann Arbor, MI: University Microfilms.
- Cooke, J. R. 1968. *Pronominal reference in Thai, Burmese, and Vietnamese*. (University of California Publications in Linguistics 52). Berkeley: University of California Press.

- Coveney, A. 2000. Vestiges of *nous* and the 1st person plural verb in informal spoken French. *Language Sciences* 22. 447–481.
- Cysouw, M. 2003. *The paradigmatic structure of person marking*. (Oxford Studies in Typology and Linguistic Theory). Oxford: Oxford University Press.
- De Rijk, R. P. 2008. *Standard Basque: A progressive grammar*. Cambridge, MA: MIT Press.
- Forchheimer, P. 1953. *The category of person in language*. Berlin: Walter de Gruyter.
- Greenberg, J. H. 1986. Introduction: Some reflections on pronominal systems. In U. Wiesemann (ed.), *Pronominal systems* (Continuum, Schriftenreihe zur Linguistik 5), xvii–xxi. Tübingen: Gunter Narr.
- Gregersen, E. A. 1974. The signalling of social distance in African languages. In W. W. Gage (ed.), *Language in its social setting*, 47–55. Washington, DC: Anthropological Society.
- Haase, M. 1992. *Sprachkontakt und Sprachwandel im Baskenland: Die Einflüsse des Gaskognischen und Französischen auf das Baskische*. Hamburg: Buske.
- Haase, M. 1994. *Respekt: die Grammatikalisierung von Höflichkeit*. (Edition Linguistik 44). Munich & Newcastle: LINCOM Europa.
- Haberl, H. & R. Hanson. n.d. Gender specification in inclusive pronouns: An analysis of Nama. Typescript.
- Haiman, J. n.d. Pronouns in Khmer. Undated field notes.
- Head, B. 1978. Respect degrees in pronominal reference. In J. H. Greenberg, C. A. Ferguson & E. Moravcsik (eds.), *Universals of human language*, Volume 3, *Word structure*, 151–211. Stanford: Stanford University Press.
- Heath, J. 2004. Person. In G. Booij, C. Lehmann, J. Mugdan & S. Skopeteas (eds.), *Morphologie—Morphology: Ein internationales Handbuch zur Flexion und Wortbildung*. (Handbücher zur Sprach- und Kommunikationswissenschaft 17. 2.), Volume 2, 998–1015. Berlin: Mouton de Gruyter.
- Heine, B. & T. Kuteva. 2007. *The genesis of grammar: A reconstruction*. (Studies in the Evolution of Language 9). Oxford: Oxford University Press.
- Heine, B., U. Claudi & F. Hünemeyer. 1991. *Grammaticalization: A conceptual framework*. Chicago: University of Chicago Press.
- Heine, B. & T. Kuteva. 2002. *World lexicon of grammaticalization*. Cambridge: Cambridge University Press.
- Heine, B. & M. Reh. 1984. *Grammaticalization and reanalysis in African languages*. Hamburg: Buske.
- Heine, B. & K.-A. Song. In press. On the grammaticalization of personal pronouns. Typescript. Chonnam National University, Gwangju.
- Held, G. 1999. Submission strategies as an expression of the ideology of politeness: Reflections on the verbalisation of social power relations. *Pragmatics* 9. 21–36.
- Hopper, P. J. & E. C. Traugott. 2003. *Grammaticalization*. Cambridge: Cambridge University Press.
- Humboldt, W. von. 1830. *Über die Verwandtschaft der Ortsadverbien mit dem Pronomen in einigen Sprachen*. Berlin: Königl. Akademie der Wissenschaften.
- Ingram, D. 1978. Typology and universals of personal pronouns. In J. H. Greenberg, C. A. Ferguson & E. Moravcsik (eds.), *Universals of human language*, Volume 3, *Word structure*, 213–247. Stanford: Stanford University Press.
- Jenkins, R. S. 2002. *Language contact and composite structures in New Ireland, Papua New Guinea*. Ann Arbor, MI: UMI Dissertation Services.
- König, E. & P. Siemund. 2000. Intensifiers and reflexives: A typological perspective. In Z. Frajzyngier & T. S. Curl (eds.), *Reflexives: Forms and functions*. (Typological Studies in Language 40), 41–74. Amsterdam: John Benjamins.

- Lakoff, G. & M. Johnson. 1980. *Metaphors we live by*. Chicago: University of Chicago Press.
- Lee, C. L. 1999. The implications of mismatched personal pronouns in Chinese. *Text* 19(3), 345–370.
- Lehmann, C. 1995. *Thoughts on grammaticalization*. Munich: Lincom Europa.
- Listen, P. 1999. *The emergence of German polite Sie: Cognitive and sociolinguistic parameters*. (Berkeley Insights in Linguistics and Semiotics 32). New York: Peter Lang.
- Lockwood, W. B. 1968. *Historical German syntax*. Oxford: Clarendon.
- Martelotta, M. E. & M. M. Cezario. In press. Grammaticalization in Brazilian Portuguese.
- Merlan, A. 2006. Grammatikalisierungstendenzen im Portugiesischen und Rumänischen: von Nominalsyntaxen zu Pronomina. In J. Schmidt-Radefeldt (ed.), *Portugiesisch kontrastiv gesehen und Anglizismen weltweit*. (Rostocker Romanistische Arbeiten 10), 221–240. Frankfurt: Peter Lang.
- Mühlhäusler, P. & R. Harré. 1990. *Pronouns and people: The linguistic construction of social and personal identity* (Language in Society 15). Oxford: Blackwell.
- Nicolle, S. In press. Pragmatic aspects of grammaticalization. In H. Narrog & B. Heine (eds.), *The Oxford handbook of grammaticalization*. Oxford: Oxford University Press.
- Pfau, R. & M. Steinbach. 2006. Modality-independent and modality-specific aspects of grammaticalization in sign languages. *Linguistics in Potsdam* 24, 5–98.
- Pfau, R. & M. Steinbach. In press. Grammaticalization in sign languages. In H. Narrog & B. Heine (eds.), *The Oxford handbook of grammaticalization*. Oxford: Oxford University Press.
- Postal, P. 1969. On so-called pronouns in English. In D. Reibel & S. Schane (eds.), *Modern studies in English*, 187–224. New Jersey: Prentice Hall.
- Samarin, W. J. 2002. Plurality and deference in urban Sango. In V. Carstens & F. Parkinson (eds.), *Advances in African linguistics* (Trends in African Linguistics 4), 299–311. Trenton, NJ & Asmara: Africa World Press.
- Shibatani, M. 1985. Passive and related constructions. *Language* 61, 821–848.
- Siewierska, A. 2004. *Person*. Cambridge: Cambridge University Press.
- Simon, H. J. 1997. Die Diachronie der deutschen Anredepronomina aus Sicht der Universalienforschung. *Sprachtypologie und Universalienforschung* 50, 267–281.
- Simon, H. J. 2003. *Für eine Grammatische Kategorie 'Respekt' im Deutschen: Synchronie, Diachronie und Typologie der deutschen Anredepronomina*. Tübingen: Max Niemeyer.
- Song, K.-A. 2002. Korean reflexives and grammaticalization: A speaker-hearer dynamic approach. *Sprachtypologie und Universalienforschung* 55(5), 340–353.
- Traugott, E. C. & R. B. Dasher. 2002. *Regularity in semantic change*. Cambridge (Studies in Linguistics 96). Cambridge: Cambridge University Press.
- Traugott, E. C. & B. Heine (eds.). 1991a. *Approaches to grammaticalization*. Volume 1. Amsterdam: John Benjamins.
- Traugott, E. C. & B. Heine (eds.). 1991b. *Approaches to grammaticalization*. Volume 2. Amsterdam: John Benjamins.
- Tucker, A. N. 1940. *The Eastern Sudanic languages*. Volume 1. New York: Oxford University Press.

