

Feature-geometry and diachrony

The development of the subject clitics in Cushitic and Romance*

Mauro Tosco

Università di Napoli "L'Orientale"

Several East and South Cushitic languages of East Africa have a preverbal series of subject markers. They are generally clitics, sometimes phonologically independent words. Like the subject clitics of many Romance varieties, these markers display characteristic restrictions: their paradigm is often incomplete, or the same morpheme may be shared by two or more persons. In this article, the subject markers of Cushitic are first compared with the Romance subject clitics, and then analyzed in the light of the feature geometry of pronominal systems (Harley & Ritter 2002b). It is argued that feature-geometric accounts are amenable to a diachronic interpretation, and that subject markers, rather than deriving directly from independent personal pronouns, arise out of the piecemeal addition of pronominal features from a minimal system. In so doing, they move along a possibly universal path of development, whose different stages are neatly exemplified in Cushitic.

Keywords: feature geometry, personal pronouns, subject clitics, Cushitic, Romance, underspecification, syncretism

* The first draft of this article was written in 2002 during my stay at the Institut für Sprachwissenschaft of the University of Cologne, which was made possible by a scholarship from the Alexander von Humboldt Foundation. The descriptive parts have been discussed at the *XX Rëscontr antèrnassional dë studi an sla lenga e la literatura piemontèisa* (Turin, May 10–11, 2003), at the *Eleventh Italian Meeting of Afroasiatic Linguistics* (Bergamo, June 5–7, 2003), and at the Round Table of *Faits de Langues* (Paris, October 17, 2003), and have been presented in Tosco (2005a, 2005b). The theoretical implications have been discussed in particular at the Round Table *Feature geometries: From phonology to syntax* (Paris, December 3, 2005).

Thanks are due to David Heap, Patricia Cabredo Hofherr, Elizabeth Ritter, Hans-Jürgen Sasse, Graziano Savà, and the anonymous referees of *Diachronica* for their comments, suggestions, and criticisms on earlier versions. Obviously, all errors and omissions are exclusively mine.

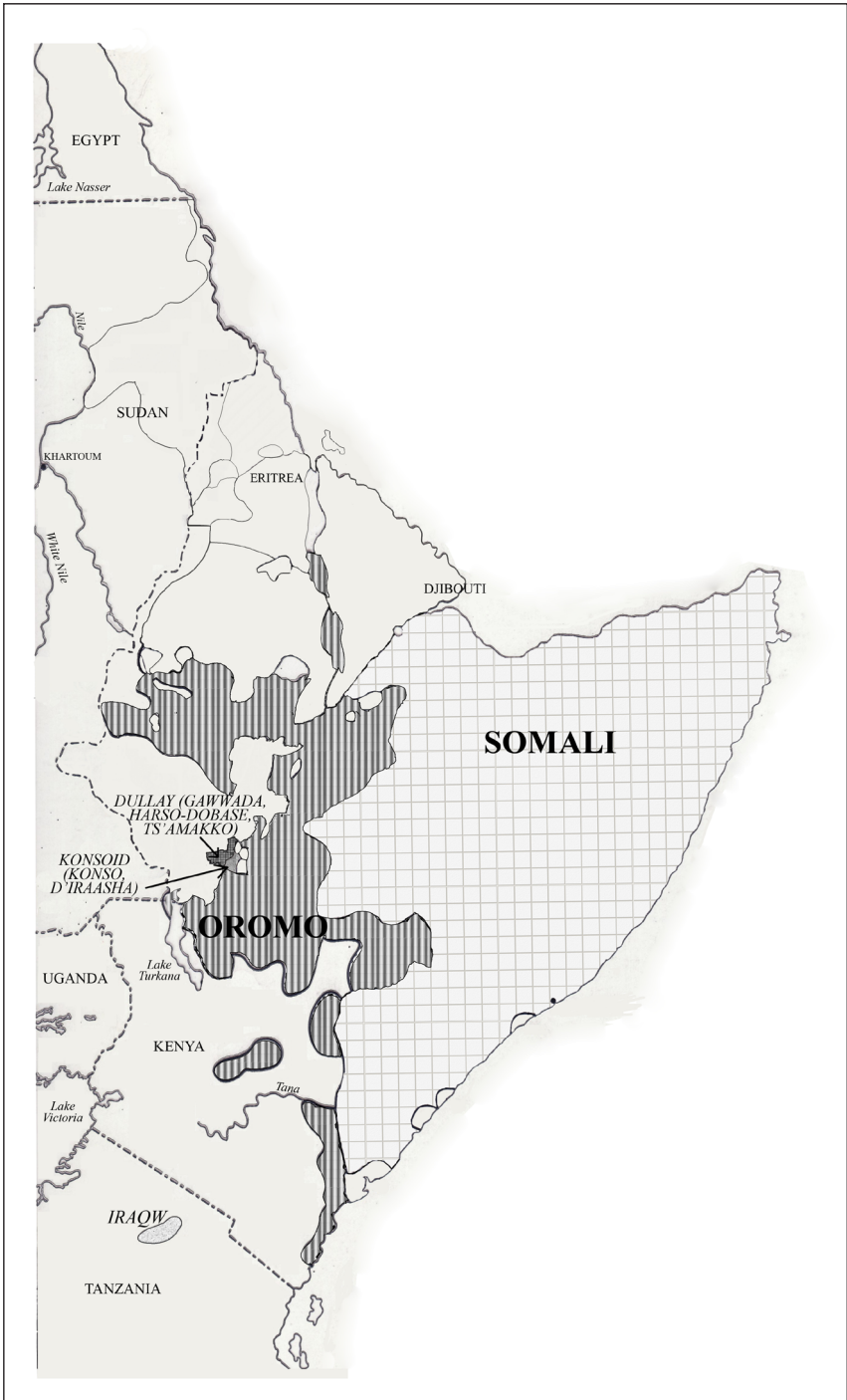
o. Introduction

Subject-Verb agreement may be expressed in languages either through independent subject pronouns, clitics (either to the verb or to a neighboring word), or straightforward verbal affixes. Historically, what is an independent word at one stage may be grammaticized and fused onto a verbal stem, passing in certain cases through an intermediate stage as a clitic. When agreement is lost or is threatened by phonetic erosion, it is often the case that a new wave of subject-marking is introduced — often through independent pronominal words, which in their turn become cliticized, and so on. On the other hand, it is also possible for such a wave of markers to arise with little or no loss of the original verbal affixes, as, e.g., when the extraposition of subject pronouns or NPs for pragmatic purposes gets over-used: In that case, through boundary shift, an original construction of the type “my ol’ man, he works all day” becomes “my ol’ man he-works all day”, and so on.

In both cases one witnesses what I shall call a Secondary Subject Marking, expressed through what I shall refer to collectively as SSMs (for “Secondary Subject Markers”).

Different from the primary (independent, clitics, or affixes) subject markers, the SSMs display characteristic restrictions in their paradigm: Typically, one finds either gaps in the paradigm of the SSMs — i.e., certain persons only (minimally, one person only) are SSM-marked — or the SSMs themselves are syncretic — i.e., one and the same morpheme is shared by two or more persons. Although both characteristics may be true of primary subject markers as well, especially when they are verbal affixes, the absence of certain markers and their syncretism seem to be on the whole much more frequent and radical among the SSMs: The latter normally do not have as many different forms as either the independent full pronouns out of which they are supposedly derived, or the verbal affixes they come to supplement. This may pose a problem for the widely held assumption that SSMs represent a “repair strategy”, since one would expect the “remedial” elements to be at least as explicit as the elements they come to rescue.

The best known and most deeply-investigated SSMs are certainly the Subject Clitics of various Romance varieties. In this article, data will be presented on the SSMs of a number of East and South Cushitic (Afroasiatic) languages of East Africa. As will become apparent, Romance Subject Clitics and Cushitic SSMs show a number of striking similarities — not so much in their syntactic behavior (which will not be considered here), nor obviously, in their phonetic substance, but in the organization of their paradigm. It will be shown that, just as in many Romance varieties, the Cushitic SSMs do NOT derive from the grafting of independent personal pronouns onto a new function. Rather, using the feature-geometric approach developed by Heidi Harley and Elizabeth Ritter (cf. in particular Harley & Ritter



Map. The East and South Cushitic languages with Secondary Subject Markers.

2002b), it will be argued that SSMs systems arise out of the piecemeal addition of pronominal features starting from a minimal system; that the initial stage always involves the expression of a single Participant (the 1st Singular in Cushitic, the 2nd Singular in Romance), followed by the expression of a single Non-Participant (the 3rd Singular), and that features involving gender and/or number may be applied to the Participants only after having been introduced for the Non-Participants. The system may of course stop at any stage, but cannot skip a step. Finally, it will be suggested that an apparent counterexample to our generalizations (the SSMs of Ts'amakko) is rather the result of the decay of SSMs (possibly due to language contact) than a stage in their development. Following up and extending Heap's (2002) analysis of the Romance Subject Clitics, feature geometry may therefore turn out to be a powerful tool in the reconstruction of tightly-knit, feature-based morphological systems.

1. Secondary Subject Markers in Cushitic: An overview

Cushitic languages generally encode subject-verb agreement through verbal affixes, either prefixed or, more commonly, suffixed. The presence of subject independent personal pronouns and/or NPs is not obligatory and is limited to "emphatic" contexts and whenever the correct identification of the subject is either impossible or difficult.

Many East and South Cushitic languages have, moreover, a series of pronominal subject markers, variously called in the literature "verbal subject pronouns", "indicators", "selectors", etc. They are different from the independent personal pronouns insofar as they are obligatory in at least a subset of the syntactic configurations of the language, and come to "supplement" the person affixes on the verbal forms. Phonologically, they are often clitics, generally proclitic to a following verbal form or an object clitic pronoun, or, more rarely, enclitic to any preceding element (as in Somali, where they follow a Focus Marker or the Declarative Classifier in main sentences, and any suitable element in subordinate clauses). They may also be autonomous words, as in South Cushitic Iraqw. Only a subset of the Cushitic languages with SSMs will be dealt with here, namely (see also the Map):

- the Central and Northern Somali dialects (East Cushitic, Omo-Tana branch), as well as the Standard Somali language based on them,
- the so-called Dullay varieties (an autonomous branch of East Cushitic),
- the Konso varieties and Oromo, which together form the Oromoid branch of East Cushitic,
- among the South Cushitic languages, the Iraqw cluster (Iraqw, Burunge, etc.).

The SSMs of a few other Cushitic languages will not be taken into account here: Rendille is on the whole quite similar to Somali; the West Omo-Tana languages Arbore and Dhaasanac (a third one, Elmolo, is extinct or almost so) have very complex pronominal systems, and the boundaries between SSMs and independent personal pronouns are still largely unclear.

The only comparative treatment of the Cushitic SSMs can be found in Robert Hetzron's influential article "The Limits of Cushitic" (1980), which deals with the SSMs only insofar as the internal classification of Cushitic (the primary objective of the article) is concerned. Hetzron called these elements "redundant subject markers", noticing how their presence is not conditioned by the presence of a nominal subject, and that they are redundant because "the verb does express person fully" (1980:67). Hetzron¹ also assumes that Somali, which (as we shall see in §2.5.) also has the richest system of SSMs, displays the original situation; the other languages would represent different stages in a process of gradual impoverishment of a maximally rich system.

While the syntax of the Somali SSMs has received a great deal of attention, especially in connection with the expression of focus, much less is known on the SSMs of other languages, and no comparative analysis has been attempted so far. This is certainly due to the ongoing absence (e.g., for much of Konsoid and Dullay) or the relatively recent availability of a grammatical description of many East and South Cushitic languages.

In certain languages, such as Somali and Iraqw, SSMs are obligatorily found in most sentence types. In Somali, this applies to declarative positive sentences both with nominal focus (marked by *baa*, to which the SSM cliticizes) and with no focus (marked by *waa*, often called a "Declarative Classifier", Saeed 1999):²

- (1) *naag baan arkay*
 woman FOC=1³ see:PST:1s
 "I saw *a woman*" (Somali)⁴

1. "This important feature of Somali syntax may be attributed to proto-SLC" (Hetzron 1980: 67; "SLC" stands for "Southern Lowland Cushitic", a genetic grouping within East Cushitic which is not accepted nowadays).

2. According to another view, *waa* is rather a marker of verbal focus (cf. Svolacchia, Mereu & Puglielli 1995). The problem is immaterial here.

3. The SSMs will be glossed with the tag of the relevant category only; e.g., in (1), the Somali 1st person (both singular and plural) SSM =*aan* is simply glossed "1".

4. Somali examples are given in the standard orthography, in use since 1972. Its main peculiarities are: <c> = [ʕ]; <x> = [ħ]; <dh> = [d]. The following abbreviations are used in the glosses:

vs.

- (1') * *naag baa arkay*
woman FOC see:PST:1s
- (2) *waan tagay*
DECL=1 go:PST:1s
“I went” (Somali)

vs.

- (2') * *waa tagay*
DECL go:PST:1s

The only instance in which SSM-marking is syntactically (although not pragmatically, cf. Gebert 1986 and Tosco 2003) optional is with a 3rd Singular Masculine subject. In (3) *waa* may be followed or not by the Subject Pronoun (i.e., the SSM) *-uu* and assume the form *wuu*:

- (3) *waa / wuu yimi*
DECL/DECL=3M come:PST:3M
“he came” (Somali)

The syntactic and pragmatic rules governing the use of SSMs in other languages are far less clear. In Gawwada (a Dullay variety),⁵ for instance, both (4) — without a SSM — and (4') — with a SSM — are grammatical:

ART	article	PF	perfect
CAUS	causative extension	PRS	present
DECL	declarative classifier	PST	past
F	feminine	S	singular
FOC	focus marker; subject-focus verbal form	SENT	sentence marker
FUT	future	1	first person
IMP	impersonal subject	2	second person
LOC	locative case (also genitive)	3	third person
M	masculine	=	clitic boundary
P	plural	.	morpheme boundary (in glosses)

5. The Gawwada data have been collected by the author in Arba Minch and in Gawwada town since 2000. I gratefully acknowledge the financial support of the Italian Ministry of Education, University and Research (M.I.U.R.) for the funding of my research, and the help of the Institute of Ethiopian Studies at Addis Ababa University for granting me the permission to carry on field-work in Ethiopia. The transcription is phonological and follows IPA conventions, except for /ʒ/ = IPA [j] and /y/ = IPA [j].

The Dullay-speakers have no overall self-denomination, nor do they seem to recognize themselves as an ethnic or linguistic unit. In Ethiopia, “Gawwada” is often found nowadays as

- (4) *ise qayna ṣande ṣúkintí*
 she tomorrow water drink:FUT-PF:3F
 “she’ll drink water tomorrow” (Gawwada)
- (4') *ise qayna ṣande i=ṣúkintí*
 she tomorrow water 3=drink: FUT.PF:3F
 “she’ll drink water tomorrow” (Gawwada)

Generally, there is evidence that the SSMs are used in Gawwada as a noun-topicalization device, and are more common with longer (modified) than with shorter, simple NPs as subjects or objects:

- (5) [*ano*] [*pako kawwádoy*] *an=ho=ṣálladeesí*
 I mouth Gawwada:M.LOC 1=you.S.M=speak:CAUS:PF:1s
 “I made you speak Gawwada” (Gawwada)

As a general rule, in Cushitic the SSMs are always excluded when the subject is in focus, which further requires suspension or reduction of subject-verb agreement. The details are different from language to language: In Gawwada the verb invariably appears in the form of the 3M singular (but with a different accentual template). In Somali there is a special accentual template, but 3F and 1P subjects still command agreement with the verbal form, the 3M form being used for all the other persons. The following two sentences from, respectively, Somali and Gawwada, exemplify the general pattern:

- (6) *nimanka baa yimi*
 men=M.ART FOC come:PST:3M:FOC
 “the men came” (Somali)

vs.:

- (6') **nimanku bay yimideen*
 men=M.ART.SUBJ FOC=3P come:PST:3P
 “the men came” (Somali)
- (7) *šéette yíʔi*
 girl eat: PF:3M:FOC
 “the girl ate” (Gawwada)

a cover term for all the Dullay-speaking groups except the Ts'amakko. In this article it will be reserved for the dialect spoken in the town of Gawwada and in the neighboring villages, while as an overall denomination “Dullay” will be used; the term “Dullay” (from the name of the most important river of the area) was originally proposed by Amborn, Minker & Sasse (1980) and has gained acceptance in the literature.

vs.:

- (7') *šéette i=yíʔti*
 girl 3=eat: PF:3F
 “the girl ate” (Gawwada)

Rather than focusing on the comparative syntax of the SSMs in Cushitic or on their etymology, as Hetzron (1980) did, this article will concentrate on the paradigms of the SSMs in Cushitic and the conditions governing their syncretism and the presence of null SSMs.

2. Syncretism and null Secondary Subject Markers in Cushitic

2.1. Oromo

Both the presence of incomplete paradigms and the syncretism in the SSMs are found in Cushitic. Syncretism is actually the norm in Cushitic, while, within the sample under analysis, the presence of incomplete sets of SSMs is limited to Oromo and to Ts'amakko. Leaving a treatment of Ts'amakko to the section on the different SSM systems of the Dullay dialects, one finds in Oromo the simplest-possible SSM-system, and one whose connection with the SSMs of other Cushitic languages seems to have been left largely unnoticed. In most Oromo varieties a 1S verbal form is, at least in many syntactic configurations (e.g., positive sentences) accompanied by a nasal element, either *=n* attached to the preceding word, or its reduplicated allomorph, the phonological word *nan*. This element is absent in the Southern Oromo dialects of Southern Ethiopia and Northern Kenya (Boraana, Waata, Orma), while it is found in the varieties of Central, Western, and Eastern Ethiopia and in the emerging written standard.

The most complete treatment of the syntax of *=n/nan* that I am aware of is found in Owens' (1985) grammar of Harar (Eastern) Oromo: In this variety, according to Owens, *=n* “signals ‘I’ and in addition emphasizes the constituent it is attached to” (Owens 1985:194). It would not, therefore, be a SSM, but what could rather be called a ‘speaker-sensitive emphasis marker’:

- (8) *magalá=n deeme*
 market=I go:PST:1s
 “I went to the market” (Owens 1985:194)⁶

6. Here and elsewhere, the transcription and glosses of the original sources are preserved.

- (9) *an isá=n arke*
 I he=I see:PST:1s
 “I saw *him*” (Owens 1985:194)

Owens underlines “to the market” and “him” in these examples in order to express “thematic emphasis”. Nevertheless, the claim that *=n* is an emphasis marker seems contradicted by Owens’ characterization of this element as obligatory with a preceding object, as in (8) and (9): How can an obligatory element express focus? What is optional is rather the full pronoun *an*. Contrast (9), which contains *an*, with (10), which lacks it:

- (10) *binensá=n ajjeese*
 animal=I kill:PST:1s
 “I killed an *animal*” (Owens 1985:195)

Moreover, when no object is found, *=n* attaches to the element *ní*, which Owens calls “a verbal focus marker”:

- (11) *an ní=n deema*
 I FOC=I see:PRS:1s
 “I am going” (Owens 1985:195)

It seems strange for an emphasizing element to combine with a verbal focus marker. On the other hand, in a few cases *=n* does apparently possess emphatic value, cf. the following:

- (12) *isá=n eerúu k’occisiise*
 he=I field till:CAUS:PST:1s
 “I made *him* till the field” (Owens 1985:195)
- (12’) *isá eerú=n k’occisiise*
 he field=I till:CAUS:PST:1s
 “I made him till *the field*” (Owens 1985:195)

How can one account for this opposition? If one assumes that *=n* is an SSM, it seems also reasonable to assume that it must immediately precede the verb, or at least not be separated from it by any autonomous element, such as an NP. For example, in the most well-studied East Cushitic language, Somali, subject and object clitics, adpositions, and other material (such as directional adverbs) relevant for the grammatical interpretation of the sentence are found within the “Verbal Piece” or “Verbal Complex”: A strictly ordered string of elements ending with the verb. Any NP lies normally outside of it. Whenever an NP does not meet this requirement and is positioned within the Verbal Piece, immediately before the verb, it is syntactically (sometimes also phonologically) incorporated with the verb and,

pragmatically, detopicalized: It will be non-referential, and it will make the first element of a noun-verb compound (cf. Tosco 2003, 2004).

Therefore, if Eastern Oromo *=n* is an SSM, (12) would actually mean something like “I made him field-till”, while (12’) would rather correspond to “the field, I made him till it”; in other words, the emphasis on the element preceding the SSM would stem from its being topical, and therefore not incorporated into the verbal form. In short:

- a. there is no emphatic marker *=n*,
- b. *=n* is a *bona fide* SSM,
- c. as such, it precedes the verbal form and follows any eventual NP,
- d. finally, the opposition in pragmatic status between an NP preceding and one following *=n* is not a matter of emphasis, but of topicality, with an SSM-following, V-preceding NP being detopicalized and incorporated.

2.2 Iraqw

SSMs have traditionally been called “selectors” in Iraqw, the southernmost Cushitic (and Afroasiatic) language. They are phonologically independent elements, and combine with Mood, Tense, Case, and affixes, yielding a remarkably complex system. They always precede the verb:

- (13) *inós i xa-xээр* *dí-r* *doo-dá’*
 3.SG 3 HAB-come:3.SG.F place:CON-F house-DEM4
 “She comes to that house” (Mous 1993:134)⁷

The SSMs are analyzed as forms of a copula verb in Mous’ (1993) grammar. But the status of *i* as a SSM is shown clearly in (13): The use of an independent pronoun (here *inós* “he/she”) is possible; moreover, the verbal form (*xa-xээр*) is marked *inter alia* for the person, number and gender of the subject. A basic conjugation in the present affirmative is found in Table 1 below for the verb *doohl*⁸ “to cultivate” (verb subclass IIIe, the largest one, Mous 1993:155):

7. The Iraqw orthography and Mous’ (1993) glosses are followed, except for the SSM *i*, which is simply glossed “3” according to our system. “CON” stands for “construct case”, “DEM” for “demonstrative”, and “HAB” for “Habitual”.

8. /hl/ stands for a lateral fricative (IPA [ɬ]) in Iraqw orthography.

Table 1. Secondary Subject Markers and Verb Structure in Iraqw

	SSMs	Verbal Structure, Present Affirmative	“to cultivate”
1S	<i>a</i>	XʋvX	<i>a dóohl</i>
2S	<i>a</i>	XʋX	<i>a dóhl</i>
3M	<i>i</i>	XʋvX	<i>i dóohl</i>
3F	<i>i</i>	XʋX	<i>i dóhl</i>
1P	<i>a</i>	XVVXáan	<i>a doohláan</i>
2P	<i>a</i>	XVVXáʔ	<i>a dohláʔ</i>
3P.M	<i>i</i>	XVVXiyáʔ	<i>i doohliyáʔ</i>
3P.F	<i>i</i>	XVVXír	<i>i doohlír</i>

The SSMs are strongly syncretic, with only two different forms, and the whole system may be represented as:

Table 2. Secondary Subject Markers in Iraqw/2

	Iraqw SSMs
1, 2 (: Participant)	<i>a</i> =
3 (: Non-Participant)	<i>i</i> =

Different SSMs but the same opposition pattern are found in the Indicative Past; the SSMs used with the Subjunctive and the Jussive display instead a person-based system (just as the one we will find in the Konsoid and Dullay varieties), with different forms for 1S and 2S:

Table 3. Secondary Subject Markers in Iraqw/3

	SSMs, Indicative Present	SSMs, Indicative Past	SSMs, Subjunctive Present
1S	<i>a</i>	<i>aga</i>	<i>ni</i>
2S	<i>a</i>	<i>aga</i>	<i>ta</i>
3M	<i>i</i>	<i>aa</i>	<i>i</i>
3F	<i>i</i>	<i>aa</i>	<i>i</i>
1P	<i>a</i>	<i>aga</i>	<i>ta</i>
2P	<i>a</i>	<i>aga</i>	<i>ta</i>
3P.M	<i>i</i>	<i>aa</i>	<i>i</i>
3P.F	<i>i</i>	<i>aa</i>	<i>i</i>

Imperative sentences, as well as Participles and Infinitives, have no SSMs. The distinction between Participant and Non-Participant is also maintained when the SSMs combine with the Object pronouns.

2.3 Konsoid

Within a restricted area of Southwest Ethiopia, the so-called Konsoid languages (belonging to the Oromoid branch of East Cushitic) represent a further step in the progressive elaboration of the system of SSMS. As far as it is known, all the varieties share the same template, with a three-fold split according to the person of the subject. The data relative to Konso, the demographically dominant variety (Black 1973), and the language spoken in and around the town of Gidole, D’iraasha or Dirayta (Hetzron 1980: 68), are shown here:

Table 4. Secondary Subject Markers in Konso and D’iraasha

	Konso SSMs	D’iraasha SSMs
1	<i>in=</i>	<i>heN=</i>
2	<i>iC=</i>	<i>het=</i>
3	<i>i=</i>	<i>he=</i>

(where *iC=* of the 2nd person in Konso copies the first consonant of the following verbal stem)

2.4 Dullay

2.4.1 *Gawwada*

Somewhat intermingled with Konsoid lies the Dullay cluster (an independent branch of East Cushitic). One finds in Dullay three different systems. In the central dialect, Gawwada, we find a first addition to the three-elements system of Konsoid: To the person-based three-fold system, an Impersonal subject clitic is added. Gawwada also exemplifies the common fact that, even when all the persons are SSM-marked, no language has as many SSMs as independent pronouns and verbal affixes:

Table 5. Independent Pronouns, Secondary Subject Markers, and Verbal Affixes in Gawwada

	Independent Personal Pro.s	SSMs	Verbal Affixes, Perfect Positive	“to drink”
1S	<i>ano</i>	<i>an=</i>	<i>-i</i>	<i>an=ʃúkí</i>
2S	<i>ato</i>	<i>aC=</i>	<i>-ti</i>	<i>aʃ=ʃúktí</i>
3M	<i>iso</i>	<i>i=</i>	<i>-i</i>	<i>i=ʃúkí</i>
3F	<i>ise</i>	<i>i=</i>	<i>-ti</i>	<i>i=ʃúktí</i>
1P	<i>ine</i>	<i>an=</i>	<i>-ni</i>	<i>an=ʃúkní</i>
2P	<i>hune</i>	<i>aC=</i>	<i>-te</i>	<i>aʃ=ʃúktí</i>
3P	<i>usunde</i>	<i>i=</i>	<i>-e</i>	<i>i=ʃúké</i>
IMP	—	<i>a=</i>	<i>-i</i>	<i>i=ʃúkí</i>

(again, *aC=* of the 2nd person copies the first consonant of the following verbal stem).

As the table shows, seven different independent personal pronouns are distinguished and the verbal form itself displays the typical Cushitic “interlocking pattern”, whereby the 1S and 3M, as well as the 2S and 3F, are identical. All in all, five different verbal forms are distinguished. In contrast, only four different SSMs are found, one for each person plus the Impersonal subject clitic (corresponding in meaning to French *on* or German *man*). The overall paradigm of the SSMs can be represented as follows:

Table 6. Secondary Subject Markers in Gawwada

1	<i>an=</i>
2	<i>aC=</i>
3	<i>i=</i>
IMP	<i>a=</i>

The similarity, even at the level of the phonological shape of the exponents, between the Gawwada (and Dullay) SSMs and the Konsoid ones seen in §2.3 is striking. Impressionistically, one can imagine the origin of the 1Sg SSM *an=* from the independent pronoun *ano* and of the 2S SSM *aC=* from *ato* (although synchronically there is no assimilation of /t/ to a following consonant in Gawwada). No obvious origin comes to mind for the plural forms in general, nor for the 3rd person SSM *i=*. Whereas for the 1st and 2nd Plural SSMs an extension of the Singular forms may be assumed, the 3rd SSM *i=* and the IMP SSM *a=* have no clear etymology.

2.4.2 Harso-Dobase

In the Harso-Dobase dialects, spoken to the east and north of Gawwada, and which form the basis of Amborn, Minker & Sasse’s (1980) description of Dullay, we find another, different evolution of the person-based three-elements system of Konsoid. As shown in Table 7, a Non-Participant (3rd person) subject is split between a form used with a 3M subject and one used for either a 3F or a 3P. While for the latter the same form used in Gawwada is used, i.e. *i=*, 3M gets a different, ‘new’ marking *u=*. This is possibly derived from the 3M Independent pronoun, which in Harso-Dobase is *uso* ‘he’, rather than *iso* as in Gawwada. Consequently, the same SSM *i=* which in Gawwada simply marks any 3rd person subject is reserved for either a 3F or a 3P subject, or, in other words, any non-Masculine subject of 3rd person.

Table 7. Secondary Subject Markers in Harso-Dobase

	Harso-Dobase (East Dullay) SSMs
1	<i>an=</i>
2	<i>aC=</i>
3M	<i>u=</i>
3F/P	<i>i=</i>

Hetzron did not know about the IMP SSM *a=* of Gawwada.⁹ Working on the basis of a three-elements system for Gawwada, he considered the Harso-Dobase system older than the Gawwada one, in line with his hypothesis of a general process of reduction of the SSMs in East Cushitic.

2.4.3 *Ts'amakko*

To the west of Gawwada a very different situation is found in Ts'amakko: In this variety, the 1st person only, both Singular and, apparently more rarely, Plural is SSM-marked. The phonological shape of the SSM is *n=*. The syntactic conditioning of the use of the SSM is still unclear; as elsewhere, SSMs and verb agreement are excluded when the subject is in focus. Two illustrative sentences are (14) and (15):

- (14) *bogolte ka mala n=basinini?*
 queen SENT how 1=do:FUT-PF:1P¹⁰
 "what shall we do with the queen?" (Savà 2005: 125)

- (15) *žiinka=ma n=zeyi*
 J.=to 1=go:PF:1s
 "I went to Jinka" (Savà 2005: 125)

As to the relationship of this system with the Gawwada one seen above, in his comparative sketch of Ts'amakko Hayward (1989:39) provides the following examples of out-of-context SSM-marked verbal forms:

- (16) *ano(m) bog'i* (i.e.: *ano (m)=boq'i*)
 I=(1) kill:PF:1s
 "I killed" (Hayward 1989: 39)
- (17) *ine(m) bog'ne* (i.e.: *ine (m)=boq'ne*)
 we=(1) kill:PF:1P
 "we killed" (Hayward 1989: 39)

(note that /*n*/ assimilates to the following consonant, becoming [*m*]). He adds, "in stark contrast to the rest of Dullay, S'amakko [: Ts'amakko, MT] appears to have shed this entire system [of SSMs, MT]. Just occasionally, a trace of the first-person selector appears." Actually, Hayward does not offer any argument in support of the idea that the simple, one-member system of Ts'amakko is just an impoverished

9. I discovered this in 2004 during my fieldwork.

10. While Savà's transcription is retained, the glosses follow those used for neighboring (and rather similar) Gawwada. An exception is made for the element *ka*, which Savà glosses "SENT" (for "Sentential marker") and which is retained as such in the glosses.

remnant of a richer system *à la* Gawwada: Nothing is known about the history of the language and the area. In principle, it could well be the other way around, i.e., Ts'amakko could preserve an earlier, less elaborated system which underwent further development in the other Dullay varieties. A look at the ethnic and cultural picture, as it often happens, may support conflicting views: On one side, the Ts'amakko are the only Dullay-speakers to live across the Weyt'o river, in close contact and alliance with the Banna, who speak a dialect of the Hamar-Banna cluster, a South Omotic language. It is probable that the impact of their Omotic neighbors has affected the language of the Ts'amakko as much as it has shaped their culture. But also the Dullay peoples to the east of Ts'amakko (the Gawwada and Harso-Dobase) are obviously not immune from contact: In particular, they are full-right members of a cultural and linguistic area which takes in the much more numerous Konsoid-speaking peoples, as well as, further to the east and north, the Burji (Highland East Cushitic) and, to some extent, the Omotic-speaking Koyra. Linguistically, the features of this "South-West Ethiopian" language area have been established by Sasse (1986). Within this area, Dullay and Konsoid occupy a central position, where convergence ranges from phonology (with, e.g., the absence of voice opposition among the stop consonants) to morphology, syntax, and lexicon (for a general treatment of the Konso-Dullay convergences in phonology, cf. Amborn, Minker & Sasse 1980: 58ff.). As far as the SSMs are concerned, it was just seen above that the similarity extends to the very shape of the SSMs and their morphophonemics, e.g., both in Konso and in Dullay the 2nd person SSM involves the copying of the first consonant of the verbal stem. Culturally and economically, the Dullay speakers seem on the whole to be in a subordinate position with respect to the Konso. In principle, the Dullay could well be on the receiving side of the linguistic influence, too.

In short, there are pros and cons to both hypotheses: That Ts'amakko is preserving a more original system of SSMs than Gawwada and Harso-Dobase, or, *à la* Hayward, that it has shed it almost completely.

At this stage, we leave the matter unresolved. As we shall see, even in the absence of historical records, certain formal clues in the shape of the paradigms of the SSMs come to our rescue and help us in confirming Hayward's intuition, namely, that what we witness now in Ts'amakko are the poor remnants of what was once, and still is further to the east, a more complex system of SSMs.

2.5 Somali

Northern-Central and Standard Somali display the maximal differentiation, i.e., the minimum amount of syncretism, in their paradigm of SSMs; Hetzron (1980: 67) considered Somali to have preserved here "the most archaic, least adulterated situation".

As anticipated, the Somali SSMs never attach proclitically to the verbal form: They are rather enclitic to a preceding element, generally (in main sentences), a Focus or Declarative Marker. They cliticize to any preceding element in Subordinate clauses. Their presence is obligatory in all declarative main sentences, except when the subject is in focus.

Number Neutralization is optional in the 1st and 2nd (participant) persons, while for a Non-Participant subject the same template characteristic of Harso-Do-base (M vs. F/P) is found. When not neutralized in the 1S form, 1P shows a difference between Exclusive and Inclusive which parallels the same distinction in the Independent Pronouns.

Table 8. Independent Pronouns, Secondary Subject Markers, and Verbal Affixes in Somali

	Independent Pro.s	SSMs	Verbal Affixes, Past Positive	“to take, bring”
1S	<i>aniga</i>	= <i>aan</i>	- <i>ay</i>	= <i>aan keenay</i>
2S	<i>adiga</i>	= <i>aad</i>	- <i>tay</i>	= <i>aad keentay</i>
3M	<i>isaga</i>	= <i>uu</i>	- <i>ay</i>	= <i>uu keenay</i>
3F	<i>iyada</i>	= <i>ay</i>	- <i>tay</i>	= <i>ay keentay</i>
1P/Excl	<i>annaga</i>	= <i>aan</i> ~ = <i>aannu</i>	- <i>nay</i>	= <i>aan</i> ~ = <i>aannu keennay</i>
1P/Incl	<i>innaga</i>	= <i>aan</i> ~ = <i>aynu</i>	- <i>nay</i>	= <i>aan</i> ~ = <i>aynu keennay</i>
2P	<i>idinka</i>	= <i>aad</i> ~ = <i>aydin</i>	- <i>teen</i>	= <i>aad</i> ~ = <i>aydin keenteen</i>
3P	<i>iyaga</i>	= <i>ay</i>	- <i>een</i>	= <i>ay keeneen</i>

**3. Syncretism and the origin of the Secondary Subject Markers:
Evidence from Romance**

Syncretism and the presence of both null and non-null forms within the paradigm of similar elements, the Subject Clitics of many Romance varieties, have received in recent years a considerable amount of attention. It has been noted (cf. Poletto 1999, 2000, and Heap 2000, 2002) that many Romance varieties have split *pro*-Drop systems, in which the verbal forms are obligatorily preceded by a Subject Clitic in certain persons only. Different from French, with its obligatory presence of a non-null Subject Clitic, and from consistent Null Subject languages, such as Spanish or Italian, these varieties have both null and non-null Subject Clitics. The comparative analysis of most varieties of Northern Italy — and to some extent of Southern France and Central Italy as well (Renzi & Vanelli 1983, Poletto 1999, 2000, Heap 2000) — shows that different systems are synchronically present. Renzi & Vanelli (1983) have originally proposed the following generalizations:

- i. languages in which only the 2S is SSM-marked,
- ii. languages in which 2S and 3S are SSM-marked,
- iii. languages with SSMs for 2S and both 3S and 3P,
- iv. “richer” systems, in which all or almost all the persons are SSM-marked.

Although within (iv) one also finds dialects in which only 1P and 2P are missing (as in Garfagnana, Italy), or which lack 1S only (as in certain Provençal varieties of Piedmont), Renzi & Vanelli’s (1983) generalizations account for the overwhelming majority of the varieties. Systems i. to iii. may also be expressed implicatively as follows, from Poletto (2000: 38):

- a. If a variety has only one SCL [: Subject Clitic, MT], this is the second-person singular.
- b. If a variety has two SCLs, these are the second-person singular, and the third person [singular, MT].
- c. If a variety has three SCLs, these are the second-person singular and the third-person singular and plural.

Thus, Renzi & Vanelli (1983: 143) have proposed to split 2S, 3S, and 3P (systems i, ii, and iii above) from the rest (systems in iv); along the same lines, Heap (2002) splits the Subject Clitics in two blocks:

2S > 3S > 3P	Block A
1S, 1P, 2P	Block B

A few incomplete SSMs systems of Northern Italy are displayed below (3rd person SSMs often show different forms for Masculine vs. Feminine in the singular and, less often, the plural; gender-opposing forms are shown for simplicity in the same line):

Table 9. Subject Clitics paradigms in a few varieties of Northern Italy

	2 elements		3 elements		4 elements
	Milan	Genoa	Feltre	Fassa (Rhaeto-Romance)	Bergamo
1S	—	—	—	—	—
2S	<i>te</i>	<i>ti</i>	<i>tu</i>	<i>te</i>	<i>te</i>
3S	<i>el</i>	<i>u (M), a (F)</i>	<i>l</i>	<i>el (M), (a) la (F)</i>	<i>(a) l (M), (a) la (F)</i>
1P	—	—	—	—	<i>(a)m</i>
2P	—	—	—	—	—
3P	—	—	<i>i</i>	<i>i (M), (al) les (F)</i>	<i>i</i>

(Rhaeto-Romance data are from Vanelli [1997], other dialects from elicitation)

Other varieties have richer systems, in which all or almost all the persons are subject-clitic-marked. Languages of this type almost always display a good amount of syncretism.¹¹

Piedmontese is an example of this type of language: While (at least in the written variety) any tensed verbal form must be preceded by a Subject Clitic, only three different clitics are used in order to cover all the persons of the paradigm. In Table 10, the independent pronouns, the Subject Clitics, the verbal affixes of the regular verbs of the 1st Conjugation in the Indicative (Main) Present, and an exemplificatory paradigm are displayed. It can be noticed that among the verbal affixes four different morphemes cover the six inflectional persons of the paradigm, with 1S and 3P sharing the same affix *-o* (/u/)¹² and 2S and 2P the affix *-e*. At the same time, the data are in keeping with Renzi & Vanelli’s (1983: 133) proposal that Subject Clitics and verbal affixes are in an inverse relationship, so that syncretism can be present either among the Subject Clitics or among the verbal affixes, but not in the same persons. The result is that the interplay of the two systems results in an unambiguous marking of all and every person.

Table 10. Independent Pronouns, Subject Clitics and Verbal Affixes in Piedmontese

	Independent Pro.s	SCs	Verbal Affixes, Indicative Present		“to play”
1S	<i>mi</i>	<i>i</i>	<i>-o</i>		<i>i gieugo</i>
2S	<i>ti</i>	<i>it</i>	<i>-e</i>		<i>it gieughe</i>
3S	<i>chiel</i> (M), <i>chila</i> (F)	<i>a</i>	<i>-a</i>		<i>a gieuga</i>
1P	<i>nojautri</i>	<i>i</i>	<i>-oma</i>		<i>i giugoma</i>
2P	<i>vojautri</i>	<i>i</i>	<i>-e</i>		<i>i gieughe</i>
3P	<i>lor</i>	<i>a</i>	<i>-o</i>		<i>a gieugo</i>

Similar systems are found in many varieties of Emilia (as in the dialect of Bologna), and in Friulian. While the actual exponents are different, Piedmontese, the dialect of Bologna and Friulian share the same syncretism, with 1S, 1P, and 2P sharing one and the same element:

11. No syncretism is found in those Provençal varieties of Piedmont which have no Subject Clitic of 1S.

12. Piedmontese data refer to the literary variety and are written in the standard orthography.

Table 11. Subject Clitics paradigms in a few varieties of Northern Italy/2

	Bologna	Friulian (Rhaeto-Romance)
1S	<i>a</i>	<i>o</i>
2S	<i>t</i>	<i>tu</i>
3S	<i>al</i>	<i>al</i> (M), <i>e</i> (F)
1P	<i>a</i>	<i>o</i>
2P	<i>a</i>	<i>o</i>
3P	<i>i</i>	<i>a</i>

(Rhaeto-Romance data are from Vanelli [1997], Bolognese data from elicitation)

Syncretism is therefore the norm, but only among the Subject Clitics of Heap’s “Block B” — 1S, 1P, and 2P. Most commonly, these persons share one and the same element, as *i* in Piedmontese, *a* in Emilia, and *o* in Friulian (Vanelli 1997: 282). Historically, it seems that in every case it was the marker of 1S which spread to 1P and 2P (Vanelli 1987: 185).¹³

This internal development is very different from the one found, e.g., in French, whose Subject Clitics straightforwardly derive from the Nominative forms of the personal pronouns of Late Latin and no syncretism is present:¹⁴

Table 12. On the origin of underspecification and syncretism: Piedmontese vs. French

	Piedmontese		Latin		French
1S	<i>i</i>	← ? —	<i>ego</i>	→	<i>je</i>
2S	<i>it</i>	← —	<i>tu</i>	→	<i>tu</i>
3S.M	{ <i>a</i> }	← ? —	<i>ille</i>	→	<i>il</i>
3S.F			<i>illa</i>	→	<i>elle</i>
1P	<i>i</i>	→	<i>nos</i>	→	<i>nous</i>
2P	<i>i</i>	→	<i>vos</i>	→	<i>vous</i>
3P.M	{ <i>a</i> }	→	<i>illi</i>	→	<i>ils</i>
3P.F			<i>illae</i>	→	<i>elles</i>

13. Cf. also Hajek (1997: 276): “The Bolognese and Romagnol systems of clitic subject pronouns are characterized by the analogical spread of /a/ < EGO from 1sg. to 1pl. and 2pl.”

14. This of course does not mean that syncretism is unknown in the French varieties (cf. Heap 2000). The use of the Subject Clitic of 1S *je* for 1P *nous* is well known in the “patois”, and is mentioned, e.g., by Nyrop (1925: 81–82):

On trouve dans les patois *j'avons, je sommes, j'étions, je savons*, etc. Cette combinaison, qui surprend par un désaccord étrange, remonte assez haut. Palsgrave a constaté son existence dès la fin du XVe siècle dans la langue vulgaire. Au XVIe siècle elle pénètre comme tant d'autres vulgarismes dans le langage des courtisans; la plupart des grammairiens la condamnent sévèrement. A cause de son caractère extrêmement populaire elle se rencontre rarement dans la littérature ... Molière s'en sert, mais il le met dans la bouche des gens de service et des paysans ... La combinaison étudiée est encore de nos jours très répandue dans les patois.

It is therefore possible to interpret Poletto's (2000) implicational scale and Heap's (2002) divide between Blocks A and B of Subject Clitics in diachronic terms, as successive steps in the development of a richer system of SSMs. In order to do this we shall make use of a feature-geometric analysis of pronoun systems, to which we turn in the next section.

4. On the feature geometry of the Cushitic Secondary Subject Markers

4.1 Background

Interest in crosslinguistic generalizations on the clustering of morphological features has a long-standing history in linguistics, especially in typologically-oriented studies. Greenberg's (1963) generalizations on the dependence of gender on number, and within the category of number (e.g., of the dependence of the presence of a dual upon the plural number) are perhaps the best example of such orientation. Much in the spirit of phonological feature geometries, crosslinguistic generalizations on the clustering of morphological features have more recently entered generative linguistics, where they are seen as an instantiation of some aspect of Universal Grammar. Morphosyntactic features are generally grouped in unstructured bundles, possibly subject to universal hierarchies and filters which rule out incompatible gender and number combinations, as in Noyer (1997). Heidi Harley

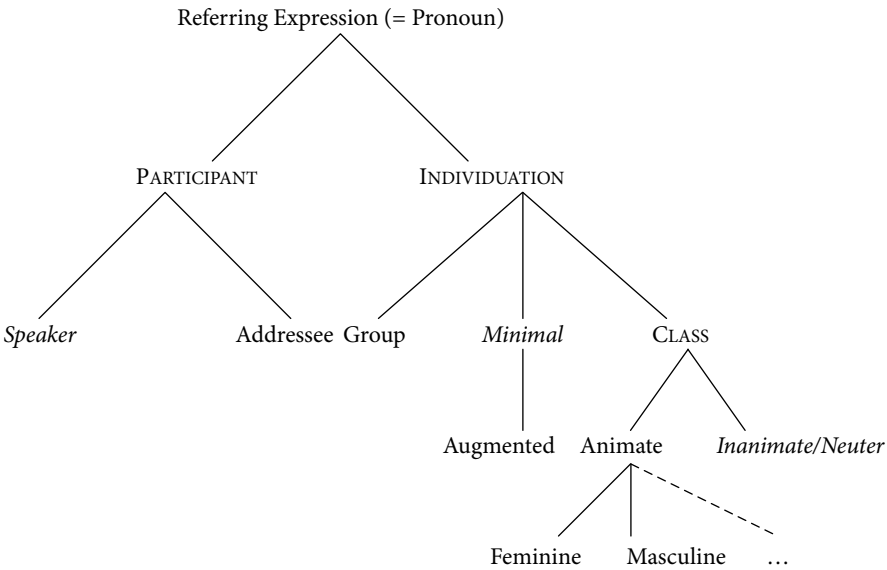


Figure 1. Harley & Ritter's (2002b: 486) feature geometry of pronominal systems

(cf. Harley 1994) and Elizabeth Ritter (Harley & Ritter 2002a, 2002b, cf. also Hanson, Harley & Ritter 2000) have recently proposed an interesting feature geometry for pronominal systems. In this proposal, features are monovalent (i.e., of the type [A] rather than [\pm A]) and only appear if they have a positive value.

The features are all dependent upon the root node RE (: Referring Expression), and are further divided into three groups, identified by the nodes in small caps: Participant (which encodes person, with the 3rd unmarked), Individuation (which represents number systems), and Class (which “encodes gender and other class information”, Harley & Ritter 2002b: 486). Within the nodes, one feature is the default value, and need not be expressed explicitly in the geometry. The default values for each node are marked in *italics*. According to Harley & Ritter, Speaker is the default value of the node Participant — a problem to which we shall return below. A 3rd person is unmarked, as it is not a Speech Act Participant: It is rather, as aptly put by Benveniste (1946), a “non-personne”. The pronouns of the familiar three-persons-two-numbers system are represented as follows (in which the default values of the single nodes are in *italics*):

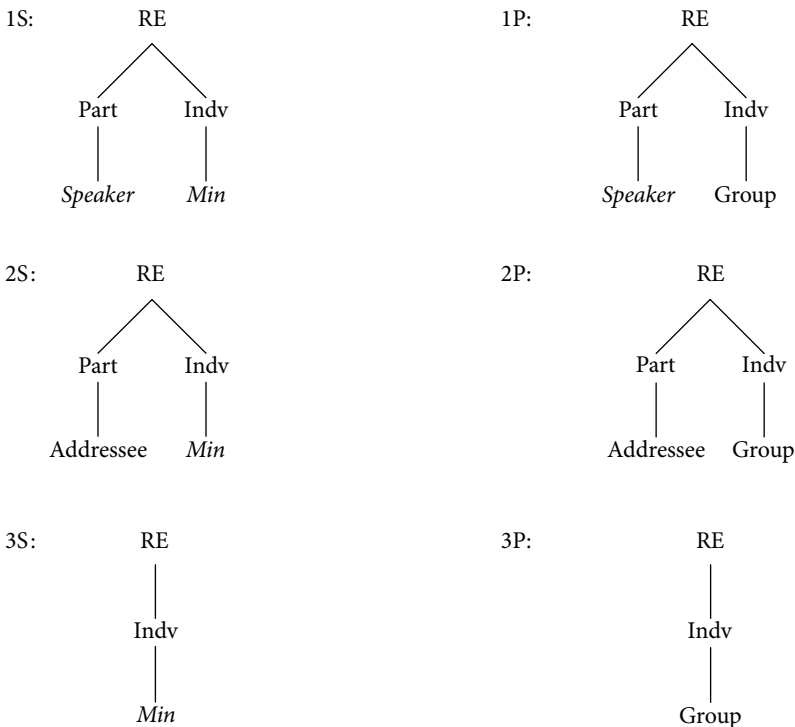


Figure 2. The feature geometry of a pronominal system with three persons and two numbers

A feature geometry of pronominal systems allows for a formal treatment of markedness: A certain combination of features will be more marked than another if it needs more nodes in order to be represented. In principle, pronominal systems can be arranged along a scale of “naturalness”, and a theoretical explanation could be provided for the absence or statistical unfrequency of highly marked systems. On the other hand, Harley & Ritter’s geometry is meant to express the combination of pronominal features operative in the overall grammar of a given language: If, e.g., a certain distinction is not made in the personal pronouns of a language but is found in verb agreement, Harley & Ritter will consider the distinction to be actually present in the pronominal system of the language, and the absence of the corresponding form among the personal pronouns to be a case of syncretism (Harley & Ritter 2002b: 510 on Koasati number, 2002b: 512 on Winnebago pronouns). This severely restricts the application of default fill-in rules: “If a node has, or can have, dependents, or if its presence is contrastive in the language, it must be represented in the underlying representation, and no default fill-in rules for that node apply” (Harley & Ritter 2002b: 498). In practice, assuming that Speaker is the default value for Part and Min the default value for Indv, they must be fully specified, as in Figure 2 above.

My aim is rather different: I take into consideration a restricted pronominal paradigm across languages rather than the whole pronominal system of a single language. I am not interested in which features are to be considered in its morphology at large. For example, a feature “Class” (or Gender) does not show up among the Subject Clitics of Piedmontese (see above, §3, Table 10), although it does in other pronominal series (the Independent pronouns and the Object pronouns). To fully specify a node found in a language but not in the paradigm under consideration would obscure the design of the single paradigm — its geometry, i.e., the way it fills up the abstract space of the pronominal expressions. I will therefore assume that, if the dependent of a node is in its turn the default value, it need not be represented. Taking 1S again as an example, as both Speaker and Minimal (=Singular) are the default values of the respective nodes Part and Indv, they can be left unmarked, and the same applies to the feature Min in the other persons of the singular, and to the feature Speaker in the 1P. The results of this radical application of default fill-in rules are shown in Table 3.

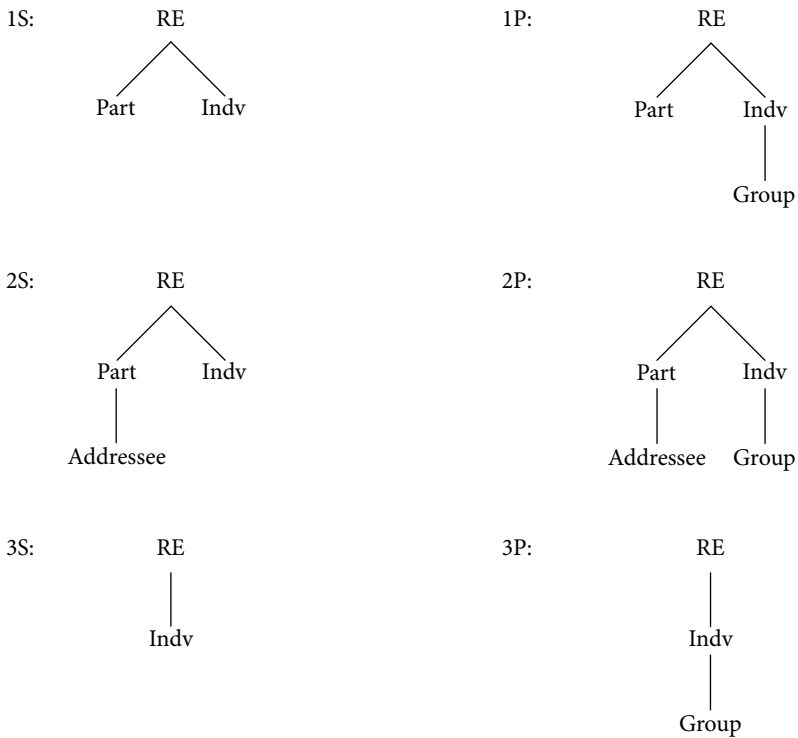


Figure 3. The feature geometry of a pronominal system with three persons and two numbers/2: the default values are not represented

Along similar lines, Heap (2002) has elaborated a feature-geometric analysis of the Romance Subject Clitics. Heap's insight is that it is possible to use feature geometry in order to account for the possible historical development of pronominal paradigms, which Heap captures through the following generalization (2002: 141):

[In] split paradigms, subject pronouns appear first among Referring Expressions which are not underlyingly specified as including Participant.

Heap (2002: 142) further predicts that "subject pronoun grammaticalisation enters grammars first via the least-marked person (or persons)". The 3rd person is unmarked (if it is really a person at all),¹⁵ but which is the "least-marked person" among 1st and 2nd? The studies on the acquisition of pronominal systems (summarized in Harley & Ritter 2002b: 499–500, cf. also Hanson, Harley & Ritter 2000), point to rather strong crosslinguistic evidence for the unmarked status of 1S

15. The logical precedence of 1S over 2S seems implied in Beneveniste's (1946) definition of 1S as "la personne subjective" and 2S as "la personne non-subjective". Both are opposed to 3, the "non-personne" (aptly called *al-yā'ibu*, "the absent one", by the Medieval Arab grammarians).

within the Part node. It is on this basis that Harley & Ritter have proposed 1S as the unmarked Participant. Other than that, nothing in linguistic theory supports the idea of the universal unmarkedness of 1S. For example, if one takes into consideration morphological markedness, one does not find 1st generally less marked than 2nd, but both 1st and 2nd marked against an unmarked 3rd (or 3S).

As was seen in §3, the basic system of the Romance Subject Clitics entails the expression of 2S only. Contrary to Harley & Ritter, Heap (2002) proposes that either Speaker or Addressee may be the unmarked value of the Participant node, and that the default value of the node Part is assigned on a language-specific basis. In Heap's analysis, therefore, 2S may group with 3S as the minimally specified element, allowing for an intuitive account of those Romance two-members systems of Subject Clitics in which 2S and 3S are expressed:¹⁶



Figure 4. A Romance Subject Clitics system {2S&3S} (modified from Heap 2002)

Also the relatively common system with Subject Clitics for 2S, 3S, and 3P is easily amenable to a feature-geometric account: 3P represents an elaboration of number ("Group"), and therefore operates within the Indv node:

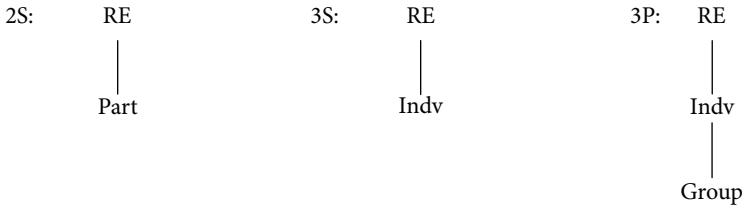


Figure 5. A Romance Subject Clitics system {2S&3S&3P} (modified from Heap 2002)

Summing up, I think it is plausible to maintain that the default value of the node Part is assigned on a language-specific basis (although I suspect that 1S is indeed on the whole less marked than 2S, and that Romance sides with a minority of the world's languages which select Addressee as the default value for Part).¹⁷

16. In Heap (2002: 138) a node Class is inserted under Indv for a Non-Participant in order to account for a Masculine vs. Feminine opposition, as in French *il* vs. *elle*. Such a distinction is not needed in those Romance varieties which have an invariable Subject Clitic of 3S.

17. The historical reasons for the "special status" of 2S in Romance are still unclear. Perhaps the inversion of the subject in the interrogative sentences — in which a 2nd person subject, especially singular, is overwhelmingly predominant — played a role, as suggested by Haiman & Benincà (1992:201–205). As noted by Renzi & Vanelli (1983:139), whenever in Romance there

Cushitic languages, on the contrary, fully adhere to Harley & Ritter's claim on the unmarked status of 1S. At the same time, the analysis of the SSM-systems of Cushitic will substantiate Heap's generalization on the development of Subject Clitic systems; it will also enable us to provide further hypotheses on the possible development of SSM-systems and the constraints they are subject to. In particular, it will be shown that SSMs systems acquire new members by adding a single feature at a time through the following steps:

- i. insertion of a bare Part node, either Speaker or Addressee, and Singular (Minimal being unmarked); this is the minimal SSMs system,
- ii. activation of the Indv node through insertion of a 3S SSM (again, the Non-Participant will be Singular, because Minimal is unmarked),
- iii. further elaboration of the Indv node through the activation of Class (e.g., Gender) and/or Number distinctions,
- iv. finally, the elaboration of the Part node through the Class/Number distinctions activated for Non-Participants in stage iii.

4.2 Geometrizing the Cushitic Secondary Subject Markers

4.2.1 *Speaker-only systems: Oromo*

The simplest Cushitic SSMs system is obviously found in Oromo, where only the 1S is SSM-marked. Keeping in mind that Speaker is the unmarked value for Part, and Min the unmarked value for Indv, and that both can therefore be left unmarked in the geometry, such a system may be represented simply as:

is inversion of the word order in the interrogative sentences and the Subject Clitic is postposed after the verb, there are at least as many different Subject Clitics as in the affirmative sentences, and often more. Also the "Block B" Subject Clitics (1S, 1P, 2P), which typically show as we have seen a characteristic syncretism in the declarative sentences, are often differentiated in the interrogatives (Renzi & Vanelli 1983:141). On the other hand, as Renzi & Vanelli (1983:134, fn. 17) point out, even in Italian, which has no Subject Clitics, there are contexts which require a 2S pronoun in order to disambiguate a sentence, while such a use is never found for the other persons: The Italian sentence *è necessario che parta subito* is only ambiguous between the meanings "it is necessary that I leave at once" and "it is necessary that he/she/it leaves at once", although the Subjunctive verbal form *parta* may be either 1S, 2S, or 3S. The point is that the sentence cannot mean "it is necessary that you leave at once", because in this case the independent personal pronoun of 2S *tu* is obligatorily present: *è necessario che tu parta subito*.

Needless to say, a similar 'preference' for 2S marking is historically present in Germanic languages, too, as shown in the development of the suffixes in *-t* in Old High German and in Middle English.

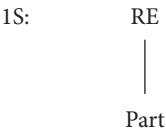


Figure 6. The feature geometry of the Oromo SSMs system {1S}

4.2.2 A ‘wrong’ system: *Ts’amakko*

Making use of Feature Geometry, we are now in a position to come back to the problem posed by *Ts’amakko* in §2.4.3. The *Ts’amakko* system, in which both 1S and 1P are SSM-marked, requires the addition of the node *Indv*, because the node *Group* (which accounts for number) is a subnode of *Indv*, but not its unmarked value (which is *Minimal*). The theory predicts that the system will develop out of *Part* by filling *Indv* before further elaborating *Part*. In other words, from a single element for Participant (1S in Cushitic, 2S in Romance) an element specified for *Indv* (3S) is added. What is excluded are systems such as $\ast\{1S, 2S\}$, $\ast\{1S, 2S, 1/2P\}$, and the like. Assuming that in Cushitic whenever both Speaker and Addressee are present, Speaker is the default value, these unlawful systems would have the following geometries:

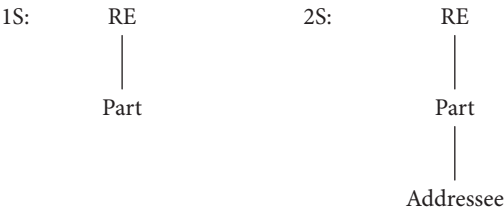


Figure 7. An impossible SSMs system: {1S&2S}

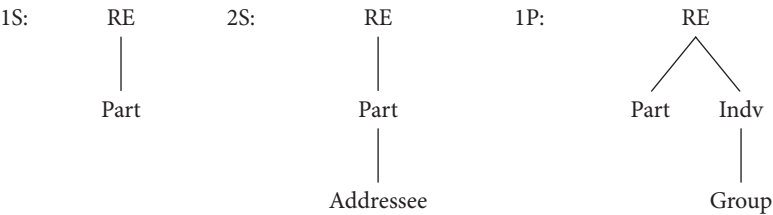


Figure 8. Another impossible SSMs system: {1S&2S&1P}

Similarly, in *Ts’amakko* both and only the 1S and the 1P are SSM-marked, while there is no SSM-marked 3rd person: This again runs against the provisions made by feature geometry, because the feature *Group*, which is found under *Individuation*, is introduced before the bare *Individuation* node. The feature geometry of the *Ts’amakko* SSMs is shown in Figure 9. The unlawful absence of the default value of *Indv* is shown by a dotted line:

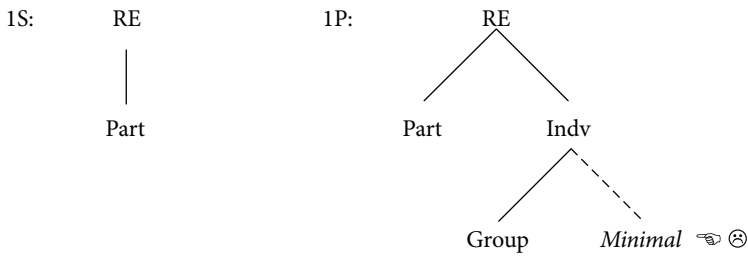


Figure 9. The Ts'amakko SSMs system {1S&1P}

4.2.3 Participant-based systems: *Iraqw*

The simplest and “lawful” expansion of the Oromo system is found in *Iraqw*, where all the persons of the paradigm are SSM-marked, but the system is highly syncretic, with a simple opposition between Participants and Non-Participants.

The feature geometry of the *Iraqw* system is shown in Table 11. The node Individuation is introduced, but no further elaborated, because no number or class distinctions are made. On the other hand, through the introduction of Individuation, the same bare node Participant which in the case of Oromo (Figure 6) coded 1S only, now codes all the Participants. This is the same pattern found in the pronominal system of Winnebago (where disambiguation is accomplished through verbal morphology), as well as in Navajo and in Lummi (Salishan, cf. Noyer 1997:112).

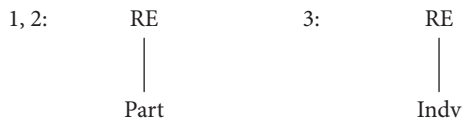


Figure 10. The *Iraqw* SSMs system {1,2&3}

4.2.4 Person-based systems: *Konsoid*

Further elaboration of the Participant node — but still with no activation of the Individuation node — is found in *Konsoid*, where each person has a different SSM, yielding a triple opposition:

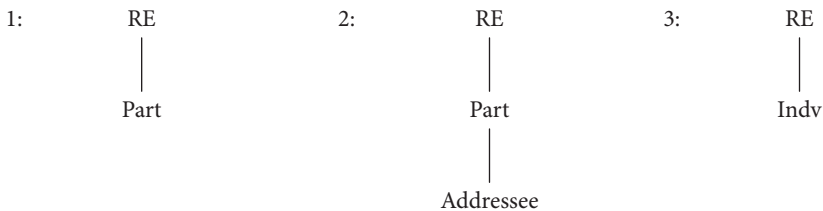


Figure 11. The *Konsoid* SSMs system {1&2&3}

Again, there is no activation of the node *Indv* (a similar analysis has been proposed by Harley & Ritter [2002b: 502] for *Pirahã*). Note that the fact that the Participant node is further elaborated with the distinction between the default Speaker and the marked Addressee (as in the “impossible systems” shown above) does not stand in contradiction with our hypothesis, because the elaboration of the *Part* node does not involve the use of features of the *Indv* node.

4.2.5 *Geometrizing the Impersonal: Gawwada*

Gawwada shows the interesting feature of a SSM for an impersonal subject. The position of the Impersonal within a feature geometry is a complicated question: The Impersonal cannot be the default value of the Individuation node — this being Minimal (i.e., a 3rd Singular). Nor can the Impersonal be put under the node Group (which identifies a 3rd Plural). Finally, the Impersonal *per se* (but see below) is obviously not an element within the Class node, which is meant to stand for the frequent subcategorization of a 3rd person according to gender, animacy, noun class, etc.

I think that a possible solution is to take the traditional label “Impersonal” seriously, and consider it to stand for, simply, an unspecified subject, either a Participant or not, without any further elaboration of the nodes *Part* or *Indv*: IMP is simply a bare Referring Expression. I argue that such an analysis does not stand in contradiction to the theory: The order of elaboration of the nodes is not violated (nothing in the theory predicts the introduction of the bare node RE) and the node *Indv* is not activated. In fact, there is no activation at all. Under such an analysis the Gawwada system could be represented as follows:

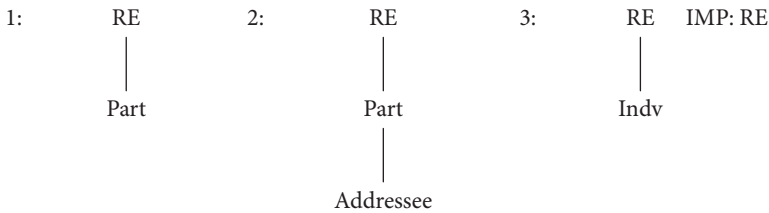


Figure 12. The Gawwada SSMs system {1&2&3&IMP}

A possible major problem with this proposal comes from the fact that an Impersonal never codes just any subject whatsoever: The subject must be an agent, or even a human. A non-agentive subject cannot be expressed by an Impersonal. In other words, impersonal constructions are basically ergative (Bruhn de Garavito, Heap & Lamarche 2003). This limitation would be lost were IMP subsumed under a bare RE node. My position is that a feature geometry of pronouns can take into account all and only those features which find a realization in the morphological

system of the language. A language with a noun class “Human” will instantiate the relevant node in the feature geometry of its pronouns (under the node Class); but the fact that the feature [\pm human] plays a role in the assignment of the subject role does not imply that the same feature is expressed in the morphology *stricto sensu*. Not differently, Bruhn de Garavito, Heap, & Lamarche (2003) have argued that French and Spanish impersonal *se* is underspecified under a node “Cl.” (for “Clitic”, equivalent to Harley & Ritter’s RE node). They add that “although *se* is underspecified morphosyntactically, it does seem to have semantic content” (Bruhn de Garavito, Heap & Lamarche 2003: 51).¹⁸

4.2.6 *Individuation-enriched systems: Harso-Dobase*

The activation of the Individuation node is the next step. In Cushitic as in Romance, this involves the elaboration of the Non-Participant. While in Romance both gender and number distinctions are apparently simultaneously introduced, in Cushitic a single feature at a time is added. Moreover, in Dullay as in many other Cushitic languages, Plural is a “third gender” alongside (Singular) Masculine and (Singular) Feminine (and pluralization therefore involves a change in gender). In both the Harso and Dobase varieties of Dullay and in Northern Somali the feature which is activated in the geometry is Masculine. In Harso and Dobase the SSM of a Masculine subject noun is *u=*, very possibly derived from the independent personal pronoun, which in these dialects is *uso* (Amborn, Minker & Sasse 1980: 97) against Gawwada *iso* (cf. Table 5). Any subject noun which is not Masculine, i.e., either (Singular) Feminine or Plural, retains the morpheme *i=* which in Gawwada marks all the Non-Participants. There is therefore no need to suppose a Generic class opposed to Masculine, as the *i=* SSM is simply assigned to whatever element is *not* Masculine. The Harso-Dobase system may be represented as follows:

18. Elizabeth Ritter (personal communication, Feb. 1, 2006) lists a few points which suggest that a feature “Human” does play a role in morphosyntax. Apart from impersonal subjects, there are, e.g., such universal or language-specific facts as person-case constraints, dative clitics in Spanish, and strong vs. weak personal pronouns in many languages. These and other issues are certainly relevant if one gives feature geometry the ambitious task of describing the set of features operative in the overall grammar of a language. In my view, the feature geometry of pronouns has a very limited goal: it is a device which describes the features which are relevant within a specific paradigm only of that language (cf. also §4.1 above).

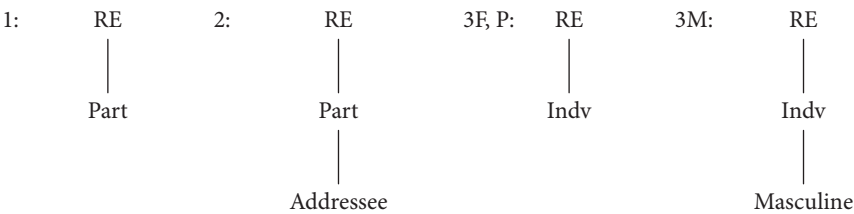


Figure 13. The Harso-Dobase SSMs system {1&2&3&3M}

4.2.7 *Optional enrichment: Somali*

The most elaborate system to be taken into examination is found in (Northern) Somali, where separate forms are found for the Plural Participants, while Indv is split, as in Harso-Dobase, along with a Masculine vs. Non-Masculine opposition. It is only at this stage that we find further elaboration of the Participant node using a feature of the Individuation node, i.e., Group. That the Participants make use of a feature (the default value Minimal vs. Group) which is not activated in the Non-Participant does not constitute, again, a violation of Heap’s generalization, as nothing prescribes the use of the same features for Participants and Non-Participants. Moreover, as mentioned earlier, while the use of the plural forms of the SSMs for the Participants is syntactically optional, such an optionality is not found for the Non-Participant. Thus, both (18) and (18’) are acceptable:

- (18) *waan tagnay*
DECL=1 go:PST:1P
“we went”
- (18’) *waannu tagnay*
DECL=1P:EXCL go: PST:1P
“we went”

On the contrary, (19) is ungrammatical, while (19’) is correct:

- (19) * *wuu yimaadeen*
DECL=3M come:PST:3P
“they came”
- (19’) *way yimaadeen*
DECL=3P come: PST:3P
“they came”

Just as in Harso-Dobase, the gender of a Plural Non-Participant (which in Somali can be Masculine or Feminine, and often is the reverse of the gender of the Singular) has no bearing on the system of SSMs. The geometry of the Somali SSMs is:

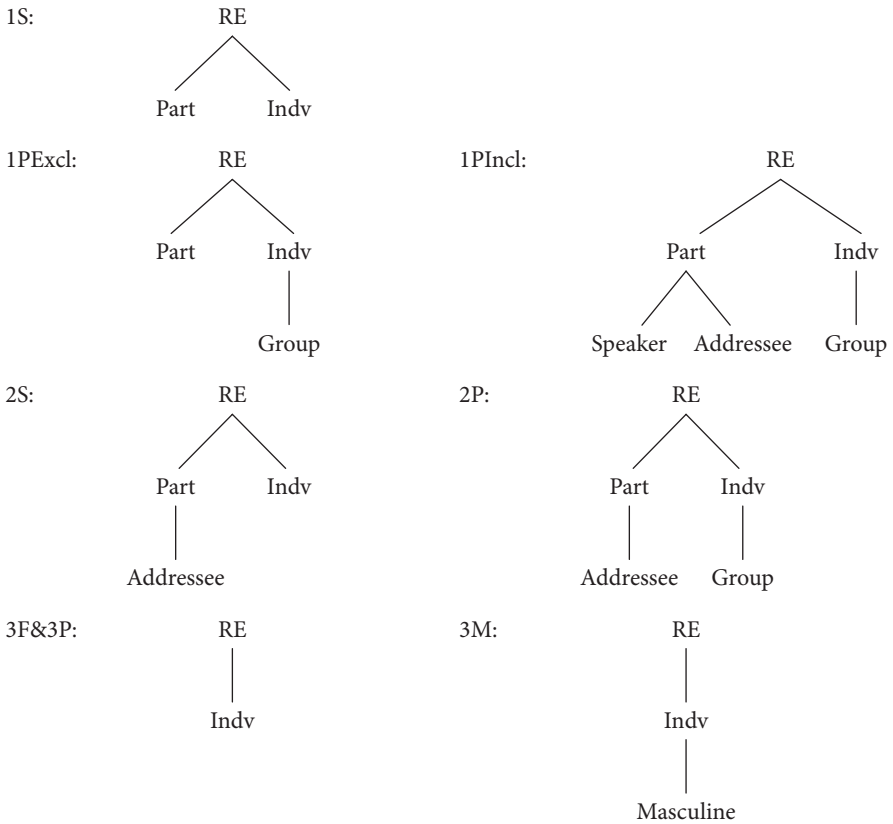


Figure 14. The (Northern) Somali SSMs system {1 (~1PExcl, 1PIncl)&2 (~2P)&3&3M}

5. Conclusions

Just as syncretism of the Subject Clitics is the rule in Romance, we have seen it to be common among the Cushitic SSMs, and along very similar lines. From the point of view of internal reconstruction, it becomes apparent that Hetzron's (1980) conclusions about Somali, with its highly developed SSM system, preserving the most archaic situation, do not hold. Quite to the contrary, the oldest SSM systems seem to be the simplest ones: Somali is younger than Harso-Dobase and Gawwada, which in their turn are younger than Iraqw. The simplest possible system is found in Oromo. What about Ts'amakko? Admittedly, any 'proof' which can be inferred from Feature Geometry is at best an indirect one, which gets its value only from the lack of direct, historical data. But the fact that it instantiates a theoretically 'impossible' system seems to be a further proof of Hayward's (1989) hypothesis that Ts'amakko has lost most of the Dullay-inherited system of SSMs, and its system of

SSMs is a poor remnant of an originally richer paradigm of SSMs, as still found in neighboring Gawwada, rather than representing the original situation of Dullay.

SSMs systems may contract and even disappear altogether. In Romance, during approximately the same span of time in which the Subject Clitics were extended to all the persons of the paradigm in a few languages, such as Piedmontese, a complete (and syncretic) similar system was apparently reduced in most varieties of Veneto, yielding the present-day system in which only 2S and both 3S and 3P are SSM-marked (Poletto 1995) — although in this case a perfectly “regular” system was the result.

Just as SSMs can spread and contract in time, they are particularly prone to areal diffusion. A glance at a map of the Romance Subject Clitics systems (as found in Heap 2000) suffices to show how very close, and very much intercomprehensible, varieties may starkly differ in the number and function of their SSMs. SSMs may be acquired quite easily; e.g., the Piedmontese system of Subject Clitics has apparently spread to many varieties of Valdôtain (belonging to a different sub-branch of Romance) in recent times (Roberts 1993). In Cushitic, the same variability is seen within Dullay: Gawwada and Ts’amakko, very similar at all levels of analysis and very much mutually comprehensible, differ, among other things, precisely in their SSMs. Within the more diversified Somali dialects one sees an even more striking difference: From the very rich SSM system of the Northern dialects to their complete absence in the Southern varieties.

A feature-geometric account can also provide an intuitive account of the fact that SSMs, although ultimately derived from independent pronouns, are not grafted directly from them: From a simpler system a richer system develops, in which, e.g., an original 1S SSM is extended to use to other persons, most typically 1P and, sometimes, 2P. Just as in many Romance varieties the Subject Clitic of 1S spread to 1P and, often but less regularly, to 2P, we find in Gawwada the extension of 1S *an=* onto 1P, of 2S *aC=* onto 2P and of 3S *i=* very possibly from 3S onto 3P. Iraqw adopted a more radical solution, extending what looks like, from a comparative point of view, a marker of 1S (*a*) to all the Participants (1S&2S&1P&2P).

References

- Amborn, Hermann, Gunter Minker & Hans-Jürgen Sasse. 1980. *Das Dullay. Materialien zu einer ostkuschitischen Sprachgruppe*. Berlin: Dietrich Reimer.
- Benveniste, Émile. 1946. “Structure des relations de personne dans le verbe”. *Bulletin de la Société de Linguistique de Paris* 43.1–12.
- Black, Paul. 1973. *Draft sketch of Konso Phonology, Morphology, and Syntax*. Manuscript.
- Bruhn de Garavito, Joyce, David Heap & Jacques Lamarche. 2003. “French and Spanish *se*: underspecified, not reflexive”. *Proceedings of the 2002 Annual Conference of the Canadian Lin-*

- guistic Association*, ed. by Sophie Burelle & Stanca Somesfalean, 42–54. Montréal: Département de linguistique et de didactique des langues, Université du Québec à Montréal.
- Gebert, Lucyna. 1986. "Focus and Word Order in Somali". *Afrikanistische Arbeitspapiere* 5.43–69.
- Greenberg, Joseph H. 1963. "Some universals of grammar with particular reference to the order of meaningful elements". *Universals of Language*, ed. by Joseph H. Greenberg, 73–113. Cambridge, MA: MIT Press.
- Haiman, John & Paola Benincà. 1992. *The Rhaeto-Romance Languages*. London: Routledge.
- Hajek, John. 1997. "Emilia-Romagna". *The Dialects of Italy*, ed. by Martin Maiden & Mair Parry, 271–278. London: Routledge.
- Hanson, Rebecca, Heidi Harley & Elizabeth Ritter. 2000. "Underspecification and Universal Defaults for Person and Number Features". *Proceedings of the 2000 Meeting of the Canadian Linguistics Association*, 11–18. (University of Toronto Working Papers in Linguistics.)
- Harley, Heidi. 1994. "Hug a Tree: Deriving the Morphosyntactic Feature Hierarchy". *Papers on Phonology and Morphology*, ed. by Andrew Carnie, Heidi Harley & Tony Bures, 289–320. (= MIT Working Papers in Linguistics, 21). Cambridge, MA: MIT Press.
- Harley, Heidi & Elizabeth Ritter. 2002a. "Structuring the Bundle: A universal morphosyntactic feature geometry". *Pronouns — Grammar and Representation*, ed. by Horst J. Simon & Heike Wiese, 23–39. Amsterdam: John Benjamins.
- Harley, Heidi & Elizabeth Ritter. 2002b. "Person and number in pronouns: A feature-geometric analysis". *Language* 78.482–526.
- Hayward, Richard J. 1989. "Notes on the Language of the S'aamakko". *Journal of Afroasiatic Languages* 2.1–53.
- Heap, David. 2000. *La variation grammaticale en géolinguistique: les pronoms sujet en roman central*. München: Lincom Europa.
- Heap, David. 2002. "Split Subject Pronoun Paradigms: Feature Geometry and Underspecification". *Current Issues in Romance Languages. Selected Papers from the 29th Linguistic Symposium on Romance Languages (LSRL), April 8–11, 1999, Ann Arbor*, ed. by Teresa Satterfield, Christina Tortora & Diana Cresti, 129–144. Amsterdam & Philadelphia: John Benjamins.
- Hetzron, Robert. 1980. "The Limits of Cushitic". *Sprache und Geschichte in Afrika* 2.7–126.
- Mous, Maarten. 1993. *A Grammar of Iraqw*. Hamburg: Helmut Buske.
- Noyer, Rolf. 1997. *Features, positions, and affixes in autonomous morphological structure*. New York: Garland.
- Nyrop, Kristoffer. 1925. *Grammaire historique de la langue française. Tome cinquième*. Copenhagen: Gyldendalske Boghandel, Nordisk Forlag.
- Owens, Jonathan. 1985. *A Grammar of Harar Oromo (Northeastern Ethiopia)*. Hamburg.
- Poletto, Cecilia. 1995. "The Diachronic Development of Subject Clitics in North Eastern Italian Dialects". *Clause Structure and Language Change*, ed. by Adrian Battye & Ian Roberts, 295–324. New York: Oxford University Press.
- Poletto, Cecilia. 1999. "The internal structure of AgrS and subject clitics in the Northern Italian dialects". *Clitics in the Languages of Europe*, ed. by Henk van Riemsdijk, 581–620. Berlin: Mouton de Gruyter.
- Poletto, Cecilia. 2000. *The higher functional field: evidence from northern Italian dialects*. New York: Oxford University Press.
- Renzi, Lorenzo & Laura Vanelli. 1983. "I pronomi soggetto in alcune varietà romanze". *Scritti linguistici in onore di Giovan Battista Pellegrini. I*, 121–145. Pisa: Pacini.
- Roberts, Ian. 1993. "The nature of subject clitics in Franco-Provençal Valdôtain". *Syntactic Theory and the Dialects of Italy* ed. by Adriana Belletti, 319–353. Torino: Rosenberg & Sellier.

- Saeed, John I. 1999. *Somali*. Amsterdam: John Benjamins.
- Sasse, Hans-Jürgen. 1986. "A Southwest Ethiopian Language Area and Its Cultural Background". *The Fergusonian Impact. Volume 1*, ed. by Joshua A. Fishman, Andrée Tabouret-Keller, Michael Clyne, Bh. Krishnamurti & Mohamed Abdulaziz, 327–342. Berlin: Mouton de Gruyter.
- Savà, Graziano. 2005. *A Grammar of Ts'amakko*. Köln: Köppe.
- Svolacchia, Marco, Lunella Mereu & Annarita Puglielli. 1995. "Aspects of discourse configurationality in Somali". *Discourse Configurational Languages* ed. by Katalin É. Kiss, 65–98. New York: Oxford University Press.
- Tosco, Mauro. 2003. "A Whole Lotta Focusin' Goin' On:' Information packaging in Somali texts". *Studies in African Linguistics* 31.27–53.
- Tosco, Mauro. 2004. "Between zero and nothing: transitivity and noun incorporation in Somali". *Studies in Language* 28.83–104.
- Tosco, Mauro. 2005a. "La naissance d'une catégorie morphologique: les clitiques sujet entre couchitique et langues romanes". *Faits de langues* 26.203–215.
- Tosco, Mauro. 2005b. "A Comparative View at the Subject Clitics in Cushitic". *Studi afroasiatici. Contributi presentati all'XI Incontro italiano di linguistica camito-semitica (Bergamo, 5–7 giugno 2003)*, ed. by Alessandro Mengozzi, 167–176. Milano: Franco Angeli.
- Vanelli, Laura. 1987. "I pronomi soggetto nei dialetti italiani settentrionali dal Medio Evo a oggi". *Medioevo Romanzo* 13.173–211.
- Vanelli, Laura. 1997. "Friuli". *The Dialects of Italy* ed. by Martin Maiden & Mair Parry, 279–285. London: Routledge.

Résumé

Maintes langues couchitiques orientales et méridionales de l'Afrique Orientale disposent d'une série de marqueurs préverbaux du sujet — habituellement clitiques, parfois des mots phonologiquement indépendants. Comme les clitiques sujet de plusieurs langues romanes, ces marqueurs présentent des restrictions typiques dans leur paradigme : celui-ci est souvent incomplet, ou alors le même marqueur est commun à deux ou plusieurs personnes. Dans l'article, les marqueurs sujet des langues couchitiques sont d'abord comparés aux clitiques sujet romans et ensuite analysés à la lumière de la géométrie des traits des systèmes pronominaux (Harley and Ritter 2002b). On propose une réinterprétation de la géométrie des traits pronominaux dans un cadre diachronique, en proposant que les clitiques sujet, en lieu de dériver directement de pronoms personnels indépendants, sont créés par l'addition successive de traits pronominaux à partir d'un système minimal et en suivant une ligne de développement peut-être universelle. Différentes langues couchitiques montrent bien les différentes étapes de ce processus historique.

Zusammenfassung

Mehrere ost- und südkuschitische Sprachen Ostafrikas haben eine preverbale Reihe von Subjektmarkern. Sie sind gewöhnlich Klitika, manchmal phonologisch unabhängige Wörter. Wie die Subjektklitika vieler romanischer Varietäten, stellen diese Marker charakteristische Einschränkungen dar.

kungen dar: Ihr Paradigma ist oft unvollständig oder das gleiche Morphem kann von zwei oder mehr Personen geteilt werden. In diesem Artikel werden die Subjektmarker des Kuschitischen zuerst mit den romanischen Subjektklitika verglichen und dann angesichts der Merkmalsgeometrie pronominaler Systeme analysiert (Harley & Ritter 2002b). Es wird argumentiert, dass merkmalsgeometrische Darstellungen einer diachronischen Interpretation zugänglich sind und dass Subjektkennezeichen durch das stückweise Hinzufügen pronominaler Merkmale aus einem minimalen System entstehen, anstatt direkt von unabhängigen persönlichen Fürwörtern abgeleitet zu werden. Dabei durchlaufen sie einen universalen Entwicklungsvorgang, dessen verschiedene Phasen im Cushitischen eingänglich veranschaulicht sind.

Author's address:

Mauro Tosco
Department of African and Arab Studies (D.S.R.A.P.A.)
Università di Napoli "L'Orientale"
piazza S. Domenico Maggiore, 12
I-80134 NAPOLI
Italy
E-mail: mtosco@unior.it, mauro.tosco@fastwebnet.it