

Motivation for copulas in equational clauses

ZYGMUNT FRAJZYNGIER, HOLLY KRECH, and ARMIK MIRZAYAN

Abstract

This study examines whether the presence of a structure exploited for a function A affects the choice of a structure for the grammaticalization of a function B; specifically, why some languages use a copula in present tense equational clauses and others do not. The use of the copula in the unmarked present tense form depends on the availability of the formal niche NP–NP, i.e., a sequence of two noun phrases without any additional marking. If NP–NP is not used for an attributive function, the equational clause may be coded by simple apposition of two noun phrases. However, if NP–NP codes an attributive function, the equational clause must be marked by some other means, such as a copula. The paper also contributes to the methodology of typological research in that we examine a typology within a geographically and genetically limited area in conjunction with a survey of unrelated languages.

Keywords: ambiguity, apposition, attribution, Chadic, compounding, copula, equational clause, focus, genitive, grammaticalization, juxtaposition, modification, possession, predication, sampling

1. Introduction: Formal niches and how they can be occupied

The aim of this study is to examine whether the presence of a structure exploited for a function A affects the choice of a structure for the grammaticalization of a function B. Such a question is relevant only when the same structure could be a plausible formal niche for the expression of different functions, and these functions minimally contain the same lexical categories.

Our claim is that a language does not allow the systematic use of the same formal niche for different functions; in other words, a language does not allow systematic ambiguity of functional interpretations. Discourse conditions are often said to resolve the functional ambiguity of formal structures. We are by no

means denying the existence of accidental, local ambiguities, triggered either by the use of homonyms or by narrowly localized ambiguities. Such ambiguities may indeed be resolved by discourse conditions. For example, such is the case of the ambiguity between the subject and object roles of relative clauses in Lakhota, Hebrew, Arabic, and Amharic. These ambiguities are limited only to the unmarked person, gender, and number (cf. M. Cohen 1970: 118).

For the purpose of this study we have selected two functions that may be competing for the same formal niche. One function is equational predication with nominal or pronominal arguments. The term "equational predication" refers to the function which "asserts that two referents are identical" (Matthews 1997: 116). The other function is the modification of one noun by another. Both functions must minimally include two noun phrases (NPs). The minimal formal niche for which these functions compete is a simple apposition of two NPs without any linking, case or other inflectional marking, and without any prosodic marking such as intonation, pauses, tones, stress. The apposition of two NPs may be differentiated by word order. Which element is modifying and which element is the head would be determined by the semantic scope of the nouns involved, as proposed in the discussion of Turkish (Section 7.9). The following is our expectation: if the formal niche NP–NP is already occupied by one function, the other function must be coded by different formal means. Hence there is a motivation for the grammaticalization of other formal means.

Assuming that a language recognizes which formal niches have been occupied by which functions, there are various coding options: (i) The NP–NP structure is occupied by the equational function, and the modifying function must be realized by a different structure; (ii) the formal niche NP–NP is occupied by an attributive function, and the predicative function must be coded by other formal means; (iii) a third choice is that the formal niche NP–NP is not exploited for any function whatsoever. We claim that the fourth choice, viz. that both functions are coded by the same means, is not possible because it would obliterate the distinction between the two functions.

With respect to coding option (i), the modifying constructions will have some marker of modification, e.g., an inflectional marker on the modifier (genitive case marker), a marker on the head (the "construct state" of Semitic languages), or adpositions.

With respect to coding option (ii), an equational clause in the present or other unmarked tense must have a marker of predication, which could be a copula, a verb, or some other means. It is not particularly important for the hypothesis what this marker is or what its grammaticalization sources are. The predication marker could derive historically from a pronoun or a deictic marker as noted already by M. Cohen (1924; cf. also D. Cohen 1984, Li & Thompson 1977, Katz 1998) or it could be a verb (cf. Heine et al. 1995).

With respect to (iii) we do not know of a language that does not exploit the formal niche NP–NP for some function, but we include this option as a theoretical possibility. Thus if a language does not exploit the formal niche NP–NP for predication, it does not necessarily mean that it must exploit it for a modifying function; likewise if a language does not exploit the NP–NP niche for a modifying function, it does not mean that it must exploit it for a predicative function. Such languages will necessarily have distinct overt coding of predicative and attributive functions. We discuss two languages that may lend themselves to such interpretations, Polish and French.

We do not claim directionality with respect to what function can occupy a given formal niche first. Throughout the history of a language, the formal niches may be occupied by different functions, albeit only those whose constituents belong to the same lexical categories. Further, we do not claim that the formal niche NP–NP must be exploited for some function. If a language exploits a formal niche for a function, it cannot use the same niche for another function. An examination of the form and function correlation can give us a picture of the relative chronology of grammaticalization for a given grammatical system at a specific time. We fully accept the possibility that a function A can replace a function B in the given formal niche. Let us consider a language whose equational clause is coded by apposition only. Such a language may use a copula to code focus in an equational clause. Such a semantically marked construction may eventually become semantically unmarked, and the copula could start to be used in unmarked equational clauses. Such developments have been frequently attested; cf. the development of the French negative marker *pas*. This scenario may explain the Hausa, Polish, and Quechua data discussed later in the paper.

By providing one more answer to the question of why some languages use a copula in present tense equational clauses and others do not, we contribute to the explanation of the neglected question why some languages have some constructions while others do not.

2. Previous studies of functions of the copula

The presence of the copula has traditionally been linked with tense; specifically there has been a widespread notion that the need to code tenses other than present requires a copula. However, Stassen (1994) demonstrates that the presence of a copula is not motivated by the overt coding of tense, aspect, and mood functions. A number of studies have additionally shown that the copula has a pragmatic function of coding focus (Li & Thompson 1977, Admoni 1983, and Lekant 1995 for Russian, McWhorter 1994 for Swahili). Hengeveld (1990) postulates a “support role” for the copula in Mandarin, which he formulates as follows: “[its] main function is to carry those grammatical distinctions

which cannot be expressed otherwise in a given language” (Hengeveld 1990: 291). Déprez & Vinet (1992) postulate that in Haitian Creole, the copula *se* licenses a determiner phrase predicate. In other words, the copula is required to turn a nominal phrase into a predicate. Feuillet (1998) provides an excellent summary of the research aimed at some universal characterizations of when the copula is and when it is not used, some of which are based on alleged semantic properties.

3. Scope and method

At least two types of equational clauses exist: (i) clauses without a separate subject phrase, which we will call for referential purposes “subjectless equational clauses”, and (ii) equational clauses with two nominal arguments, which we refer to as “equational clauses”. In subjectless equational clauses, the speaker identifies an element in the environment without explicitly naming it. One of the characteristics of such clauses is that they usually require a copula:

- (1) *nè bèn-íŋ*
COP brother-1SG
‘He is my brother.’ (Lele, East Chadic; Frajzyngier 2001)
- (2) *hlá yà*
cow COP
‘It is a cow.’ (Hdi, Central Chadic; Frajzyngier with Shay 2002)

There are also examples of equational clauses with a nominal subject and predicate without a copula:

- (3) *BOR też jest jednostką wojskową*
BOR also be.3SG.PRES unit.INST military.INST
‘BOR is also a military unit.’ (Polish, *Gazeta Wyborcza*, Dec. 27, 1999)
- (4) *Orson bohaterem, za to się należy cześć całkiem osobna*
Orson hero.INST for that REFL belong honor quite special
‘Orson is a hero, and for that he deserves quite a special honor.’
(Sources)¹

Our paper is concerned with equational clauses which have a nominal subject and a nominal predicate, as illustrated by (3) and (4) above. Moreover, we narrow the scope of our inquiry to the present or unmarked tense only, i.e., to the

1. “Sources” refers to electronic sources to Kurcz et al. (1990).

tense in which some languages have a copula and others do not. The reason for this choice is that we want to eliminate from consideration those clauses where the copula serves as tense carrier. There is, for example, no functional distinction between Russian and Polish in the past tense equational clause, where both use the copula; however there is a distinction between the two languages in the present tense, where Russian does not have a copula, but Polish, at least in its literary variety, does. We shall touch on locative and existential clauses only marginally. The reason why we will consider these other types of clauses is because in some languages they may use the same copula as equational clauses. We recognize that there may be different motivations for the presence of copulas in equational clauses, and the distinction between attributive and predicative function is only one of them. In the course of our analysis we discuss the other functions whenever they are available in the literature or whenever they have emerged from our own analyses.

4. Methodology

In order to find out whether there is a correlation between the form of the modifying construction consisting of two nouns and the equational clause consisting of two nouns, we have two types of samples. One is very broad, consisting of a representation of languages belonging to unrelated families and spoken in various geographical locations. The purpose of this limited sampling is to test the hypothesis on languages that did not serve as the source of data on which the hypothesis was built.

The other sample is much narrower, limited to languages belonging to one family and spoken in a contiguous geographical area. The purpose of this narrow sample group is to examine whether the structural correlation holds between targeted structures or whether it is determined by areal or genetic relations that hold among languages. Should the correlation be a product of common retention, we would expect it to be found in closely related languages. Should the correlation be a product of areal influence, we would expect it to be widespread over the area. Should the correlation be the result of structural constraints within a language, we would expect to find it only where the constraints exist, regardless of how closely the languages are related and regardless of their geographical proximity. For this narrow sample we have selected Chadic languages, a branch of the Afroasiatic family. The choice of Chadic rather than some other family is motivated by our familiarity with the languages in question, which allows better control over the data. We can analyze the function of each construction, and we can place this function within the larger frame of the grammatical system of the language. Moreover, because of our own field experience with Chadic languages, we know which data are elicited and which come from natural discourse. For the study of the functions of grammatical structures

and morphemes, it is important to have natural data. Because the elicited data are most often deprived of any context, it is never clear which semantic and pragmatic functions the given forms represent. We consider elicited data to be unreliable sources of information for the discussion of functions, particularly for the discussion of possible functions of copulas.

5. An illustration: A case study in English

In English, equational clauses are marked by the copula *to be*, while NP–NP apposition is used for modification; the genitive and other constructions are indicated by grammatical morphemes. The verb *to be* is the “principle copula” used with an NP subject complement; it is “the most neutral in meaning. It is also overwhelmingly the most common” (Quirk et al. 1985: 1173, 1174).

Data from the spoken American English corpus, *Switchboard* (documented in Godfrey, Holliman, & McDaniel 1992), provide examples of NP–NP equational clauses using the present tense of verb *to be*. The relevant portion is marked by roman print:

- (5) a. ... *because they're at a higher risk of exposure and blood screening is part of that.*
 b. *So, a lot of our clients are oil companies, big oil companies.*
 c. *And phone books are a large volume of annual trash.*

The apposition of two or more NPs codes modifying function. Such constructions are abundant in data from *Switchboard*:

- (6) a. *We're not being tested for drugs at all, uh, our policies and procedures manual, uh, the furthest it goes about drugs is in the miscellaneous section, or its reasons for immediate dismissal, ... it says, use of narcotics on company premises.*
 b. *If you take your TV in, a lot of these TV repair places will say, well, I'll repair your TV for a hundred dollars ...*
 c. *I think the retirement home idea's a nice idea.*

Some NP–NP constructions seem to be in the process of becoming compound nouns because of the frequency with which they occur as collocations; their new status as compounds is reflected in their stress pattern, as, e.g., in *his 'life story, a 'dish cloth, a 'Sussex man* (Quirk et al. 1983: 1330). Some noun–noun structures obtained from the *Switchboard* data seem to follow a similar development; cf. *news wires* and *death penalty* in (7):

- (7) a. *I'm overwhelmingly disappointed with the media in general except for the raw news wires.*
 b. *Do they have a death penalty in California?*

Though Quirk et al. note that the conditions invoking the different stress patterns are not entirely clear, the stress shift is connected with the degree to which a collocation is lexicalized as a compound lexical item. Compound nouns in *Switchboard* include *checkbook*, *database*, *payroll*, and *daycare*:

- (8)
- a. *I've found it's the only reliable way to keep a checkbook balanced...*
 - b. *It would be interesting to see how the database, so well, I guess we're talking about the experiment.*
 - c. *... so we are going to be starting a regular payroll deduction.*
 - d. *I, I don't feel comfortable about leaving my kids in a big daycare center.*

The final two examples include a compound noun followed by another noun (*payroll deduction* and *daycare center*), making them NP–NP constructions in which the modifying noun is a compound.

Another construction involving two NPs in English is the appositive. Appositives must normally “be identical in reference” and are marked by the prosodic change of pauses before and after the appositive, as indicated in writing by commas (Quirk et al. 1983: 1301): *Anna, my best friend, was here last night*. This type of construction is relatively rare in *Switchboard*, but in these few examples, the commas again indicate the insertion of pauses; cf. *MTV in Well, do, do you watch music television, MTV at all?*; and *the president in Our friend, the president, right now, says no new taxes*.

The genitive is coded by either the clitic *'s* or the preposition *of*, so that the following pair of examples has a similar meaning (Quirk et al. 1983: 321): *What is the ship's name?* and *What is the name of the ship?*. However, in other cases, both forms are not equivalent in meaning, and only one is appropriate (Quirk et al. 1983: 321): *John's school* but only marginally *the school of John*; *the front of the house* but not **the house's front*.

6. The copula in Chadic languages

Chadic languages belong to the Afro-Asiatic family. They are usually classified into three or four branches: West, Central, East, and Masa (Newman 1977, Jungraithmayr 1978). They are spoken in a contiguous area comprised of northern Nigeria, Niger, northern Cameroon, and southern Chad. Some Chadic languages have a copula in pragmatically neutral equational clauses while others do not. Most languages have several constructions coding modification of one noun by another.

Our selection of languages within the Chadic branch is dictated by the needs to represent all branches within Chadic and to have various types of structures represented. From the West Branch we have taken Hausa, Mupun, and Miya;

Table 1. *Summary of Chadic data*

Language	NP–NP Attributive	NP–NP Predicative	Equational Copula
Hausa	no	yes	no*
Mupun	yes	no	yes
Miya	no	yes	no
Gidar	yes	no	yes
Mina	no	yes	no
Hdi	no	yes	no
Lele	yes	no	yes
East Dangla	no	yes	no

from the Central Branch Gidar, Mina, and Hdi; and from the East Branch Lele and East Dangla. For all these languages except Hausa and Miya we have our own data. For Hausa we have used data from natural discourse texts rather than from descriptive grammars. We do not aim to reconstruct here the history of the relevant structures for Chadic, nor to conduct a comparative study of Chadic, but to examine in the selected languages the relationship between equational clauses and modifying constructions.

Table 1 summarizes the Chadic data; “no” means that the construction does not exist, and “yes” means that it does exist. The question about “Equational copula” is not whether there may be a copula in the equational clause, but rather whether the copula, if any, has the function of coding the clause as equational. This important distinction is elaborated several times in the present study, with respect to such languages as Hausa, Mandarin, and Polish. Whenever our analysis differs from the descriptions usually found in standard grammars we mark such an entry with an asterisk (*) and we justify our analysis further in the text. Note that in Table 1 we state that in Hausa the copula does not carry an equational function. This contradicts the usual analyses found in pedagogical and descriptive grammars; we address this discrepancy later in this study.

6.1. *Mupun (West Chadic)*

Mupun confirms our hypothesis in that if the modifying function involving two nouns is coded by juxtaposition with the head preceding the modifier, the equational clause requires a copula between the subject and the predicate. Examples of modifying constructions include:

- (9) a. *siwol laa*
 money child
 ‘child’s money’

- b. *miskoom abuur*
chief Abuor
'chief of Abuor'
- c. *jep miskoom abuur*
children chief Abuor
'children of the chief of Abuor'
- d. *tobeen caan*
handle hoe
'handle of a hoe'
- e. *tobeen caan nalon*
handle hoe Nalon
'a handle of Nalon's hoe' (Frajzyngier 1993: 151)

The copula in Mupun is *a*, and it is obligatory in all types of equational clauses:

- (10) a. *ngwe desə a miskoom Mupun*
man DEM COP chief Mupun
'This man is the chief of Mupun.'
- b. *dacya a siar fen*
Dacya COP friend 1SG
'Dacya is my friend.'
- c. *vəl kə vəl a feer*
two CONJ two COP four
'Two and two is four.'
- d. *səm wur a dakom*
name 3M COP Dakom
'His name is Dakom.'
- e. *səm desə a səm tul*
name DEM COP name home
'This name is a native name.'
- f. *wur a wat*
3M COP thief
'He is a thief.'
- g. *ngu wat nə a wur*
man steal DEF COP 3M
'The thief is he.' (Frajzyngier 1993: 254)

6.2. *Miya (West Chadic)*

In *Miya* the modifying construction, called "direct genitive" in Schuh (1998: 246), involves the addition of a vowel and a number of tone changes that the modifier noun undergoes under the influence of the head noun:

- (11) *mbàdə + laahə* → [*mbàdà làahə*]
 thigh jackal 'thigh of a jackal'
 (Schuh 1998: 246–247)

The equational clause consists of juxtaposition of the two NPs without any phonological changes (Schuh 1998: 316):

- (12) *Ndùwya miy-dzəhə*
 Nduwya Miya-man
 'Nduwya is a Miya man.' (Schuh 1998: 246–247)

Miya confirms our hypothesis since the modifying construction requires marking, while the equational clause structure is NP–NP juxtaposition.

6.3. *Gidar (Central Chadic)*

The structure NP–NP without any grammatical changes to any component and without any adpositions codes inalienable possession and part–whole relationships:²

- (13) a. *mà glà*
 mother house
 'first wife' (the term used by the husband, younger wives)
 b. *dák mólpa*
 middle river
 'middle of the river'
 c. *hlù dərhlíyé*
 meat hyena
 'flesh of hyena' (Frajzyngier fieldnotes)

Gidar confirms our hypothesis that if the NP–NP construction codes a modifying function, it cannot be used for the predicative function. Equational clauses must have a copula regardless of whether they have a nominal subject and predicate or the nominal predicate alone. Thus there are two types of predicative structures involving a copula. The first, NP Copula NP, is particularly relevant for our thesis:

2. Other modifying constructions, including alienable possession, use adpositions:

- (i) a. *hlù ná dərhlíyé*
 meat GEN hyena
 'meat of hyena, meat that belongs to hyena' (Frajzyngier, fieldnotes)
 b. *mámbà ná krá dərhlíyé tímé gəm háw-kə*
 story GEN dog hyena sheep CONJ goat-F
 'a story of a dog, a hyena, a sheep, and a goat' (Frajzyngier, fieldnotes)

- (14) a. *kəzà tó mól̩m-ù á-ká-n-dàýá*
 Kiza COP sibling-1SG REL-F-surpass
 'Kiza is my older sister.'
- b. *tìzì yí mól̩m-ù àn-dàýá*
 Tizi COP sibling-1SG REL-surpass
 'Tizi is my older brother.' (Frajzyngier fieldnotes)

The second structure has a predicate alone, an independent pronoun, or a noun followed by the copula. In this type of structure the copula is the only marker of predication.

- (15) a. *ndə-t tó*
 3-F COP.F
 'It is she.'
- b. *dərlíngè yí*
 hyena COP.3M
 'It is a hyena.' (Frajzyngier fieldnotes)

NPs alone, without a copula, would not be interpreted as equational clauses, a phenomenon attested in other languages, including other Chadic languages such as Hdi.

6.4. Hdi (Central Chadic)

Hdi belongs to the same branch of Chadic as Gidar, yet it has different structures in the modifying constructions and equational clauses. In Gidar the NP-NP niche codes modification, whereas in Hdi it codes predication. The types of structures in Hdi confirm our hypothesis. Modifying constructions have the form Noun *á* Noun, where *á* is a marker of a genitive construction (all data are from Frajzyngier fieldnotes; for a description of Hdi cf. Frajzyngier with Shay 2002):

- (16) a. *hlúwí-á kɾì* → [*hlúwá kɾì*]
 meat-GEN dog 'dog meat'
- b. *wì-á tghá* → [*wàtghá*]
 mouth-GEN door 'doorway'

In accordance with our hypothesis there is no copula in neutral equational clauses:

- (17) a. *m̀nd-á ráyá mbítsá*
 man-GEN hunt Mbitsa
 'Mbitsa is a hunter.'
- b. *m̀nd-á ráyá grá-dá*
 man-GEN hunt friend-1SG

'My friend is a hunter.'

Although a copula is not required when there is a nominal subject and predicate, it is required in subjectless equational clauses, just as in Gidar (see Section 6.3):

- (18) *hlá yà*
 cow COP
 'It is a cow.'

6.5. *Mina (Central Chadic)*

Like Gidar and Hdi, Mina belongs to the Central branch of Chadic. The modifying construction is characterized by several morphological markers intervening between two nouns. One is final vowel reduction on the head noun, an inflectional means. In the following example, the noun *̀nkùà* 'goat' has the final vowel reduced when it is followed by the modifier *m-fés* 'small' (*fés* 'small' is a property concept noun; all data are from Frajzyngier, Johnston, & Edwards 2002):

- (19) *̀nkù m-fés*
 goat REL-small
 'small goat'

Other morphological markers of modification are relative clause particles between the two components of the modifying construction, one of which is the genitive marker *tó*:

- (20) *tàlàn tó zàvàn-yí*
 head GEN guinea.fowl-PL
 'heads of the guinea fowl'

The only type of modifying construction coded by apposition of two nouns in Mina occurs when the modifying construction is a part of a locative adjunct with an inherently locative predicate, or a part of the locative adjunct preceded by the marker *á* which is a locative predicator.³ In such constructions the genitive particle is not allowed:

- (21) a. *kè dí dúwòn m̀dìngẁr̀zé*
 INF put back donkey
 'He put it on the back of the donkey.'

3. The locative predicator *á* behaves in the same way as a locative predicate. It is used when the clause has to have a locative adjunct, but the predicate is not inherently locative.

- b. *tíl zà á bíŋ kwíl-yî*
 leave be PREP room pot-PL
 'He went into the house of kulis.'⁴

The implications of this phenomenon for Mina are dealt with in Frajzyngier, Johnston, & Edwards (2002). The appearance of an appositional modifying construction in a prepositional phrase does not affect our hypothesis, since the environment in which the appositional structure occurs is itself marked. In all other environments the genitive marker must be used:

- (22) a. *tór láy tǎ mìtǎš*
 month time GEN hunger
 'the year of the hunger'
 b. *mbə mátsà*
 son mother.2SG
 'your younger brother'

In accordance with our hypothesis, equational clauses in Mina use the appositional construction NP–NP, characterized by a lack of vowel reduction of the first NP, in contrast with modifying constructions which are marked by vowel reduction.⁵ In the following example, the noun *ñkùá* does not have its vowel reduced:

- (23) *ñkùá fés*
 goat small
 'The goat is small.'

Equational clauses in Mina do not have the usual subject and aspect markers that occur when the predicate is a verb. In addition, equational clauses do not have 3rd person subject pronouns that are required when the predicate is a verb:⁶

4. The term *kuli* means 'deity', usually an ancestral spirit, and is symbolized by a pot where offerings are deposited.

5. In contrast to equational clauses, existential and locative clauses in Mina require appropriate verbs:

(i) *ngúl-nə dáhá wəzi-n-yí dáhá*
 husband-1SG exist children-1SG-PL exist

'My husband's here, my children are here.' (Frajzyngier, Johnston, & Edwards 2002)

6. In Mina the copula is not used in pragmatically neutral, specific interrogative clauses: *wà mí* DEM what 'what is this?'; *há ví* 2SG who 'who are you'. But the copula is used to code focus: *wà ží mí* DEM be what 'what is this?' (all examples from Frajzyngier, Johnston, & Edwards 2002)

- (24) a. *wà bàhá gálùmbá lèk lèk ɣlám tí hídò táŋ*
 DEM again story Lek Lek name GEN man DEM
 'This is a story of Lek. Lek is the name of that man.'
- b. *hídù wà jín*
 man DEM tall
 'The man is tall.'

6.6. *Lele (East Chadic)*

Lele is an East Chadic language spoken in southern Chad. The structure NP–NP, without any changes on either NP, and without any adpositions has a modifying function, with the first NP being the head and the second the modifier. (cf. Frajzyngier 2001 for a description of Lele; vowels without tone marker have mid tones):

- (25) a. *gidirè ba*
 moon man
 'masculine moon'
- b. *kùb lèle*
 language Lele
 'the Lele language'
- c. *kàrà túgú*
 people village
 'village people'
- d. *ɓodu tamá*
 monkey woman
 'female monkey'⁷(Garrigues-Cresswell & Weibegué 1981: 2–3)

Lele supports our hypothesis in the following way: the modifying function is coded by apposition of two nouns while the equational predicative function requires a marker. The marker is the copula *ne*, which is used with both singular and plural subjects:

- (26) a. *kiya dàgè ná gilkínín ne kósuŋ go lélé*
 Kiya 3PL COM Gilkinin COP term REFR Lele
 'Kiya and Gilkinin are Lele names.'
- b. *láràdí na nè kara go yà-gé*
 chameleon HYP COP people REFR say-3PL
 'Chameleon would represent people who say ...'

7. To alleviate any suspicion that *tamá* is an adjective, the following example provides evidence that it is indeed a noun: *tamá, ul dé* woman cry.IMP NEG 'woman, do not cry' (Garrigues-Cresswell & Weibegué 1981: 2–3).

- c. *gúló na nè kara go yà-gé*
 lizard HYP COP people REFR say-3PL
 ‘Lizard would represent people who say ...’ (Frajzyngier 2001)

6.7. *East Dangla (East Chadic)*

East Dangla, like Lele, belongs to the East Chadic branch, and yet it displays different structural properties. Nevertheless, the totality of these properties supports the proposed hypothesis. There are several constructions that use demonstratives, relative clause markers, and an oblique case marker to code the modifying function (all data, Erin Shay, p.c.):

- (27) a. *law ku bóor-rì*
 hair DEM hyena-OBL
 ‘the hair of the hyena’
 b. *dafàrr-ey ka mìʔilè-ì*
 wounds-3M.POSS DEM lion-OBL
 ‘the lion’s wounds’
 c. *wàdʔiin ta koo-rì*
 ‘little famine’ DEM millet-OBL
 ‘the shortage of millet’

Given the form of the modifying constructions in East Dangla, the appositional structure NP–NP is thus available for the equational predicate function:⁸

- (28) a. *kú kokinin yòme*
 2PL thieves INTENS
 ‘You are such thieves.’
 b. *èp-ik ànáy yòme*
 DEM-DEM much INTENS
 ‘That’s too much.’

6.8. *Hausa (West Chadic)*

Hausa is a West Chadic language. If one were to believe the existing descriptions, Hausa is an example of a language which does not exploit the NP–NP formal niche because both the modifying construction and equational clauses consist of two NPs and some additional markers, which are different for each type of function.

The modifying construction involving two nouns requires a genitive linker, historically derived from a demonstrative, suffixed to the head noun:

8. As in equational clauses, there is no copula in locative clauses in East Dangla: *tʔaa kó ʔtaanè*
 3F already here ‘she’s right here’ (Erin Shay, p.c.).

- (29) *zaurukà-n gidá*
 entry.porch.PL-GEN house
 'entry porches of the house' (Kraft 1963)

According to the proposed hypothesis, the equational clause could exploit the formal niche NP–NP. Yet, Hausa has a masculine/plural copula *ne* or feminine copula *ce*, depending on the gender and number of the head noun, following the Subject-Predicate sequence (Kraft 1963, Newman 2000):

- (30) *màganà, 'ai, jarì ce*
 speech indeed capital is
 'Speech is the capital [one uses to set up a business with].' (Kraft 1963: 21)

Newman (2000: 164) states: "In normal equational sentences, affirmative or negative, the stabilizer [copula] is obligatory". He then goes on to list cases when the copula must or may be omitted. Most of the examples in the discussion of equational clauses appear to be produced out of context, possibly answering the question "how do you say x".

Natural discourse data, however, indicate that Hausa is, in fact, a language that exploits the NP–NP niche to code equational clause; the copula is not used unless it carries a function other than simply coding equational predicate. Natural discourse data provide plenty of evidence that copulas are not obligatory components of equational clauses or locative clauses:

- (31) a. *'ya daya gàre shì*
 child one PREP 3M
 'He has one child.' (Kraft 63: 1)
- b. *dâ duk 'àbîn nan gurin wàzirì*
 formerly all thing DEM place.GEN chief
 'Formerly, everything favored the chief; lit., everything was the place of the chief.' (Kraft 1963: 5)
- c. *'ay, kùwa, gàskiya-r-ka*
 indeed again truth-GEN-2M
 'Indeed, you are right again; lit., it is your truth.' (Kraft 1963: 24)
- d. *'àbinci-n dòrina ciyawà dà 'yan tsìrè tsìre*
 food-GEN hippopotamus grass ASSC child sprout
na bàki-n ràfi
 GEN mouth-GEN water
 'The food of hippopotamus is grass and small sprouts from the edge of streams.' (Kraft 1963: 85)

However, the copula is used in a number of constructions coding a function

other than equational predication. Thus, for example, it is used in focus constructions:

- (32) *kai nè, bà kà sanì ba*
 2M COP NEG 2M know NEG
 'It is you who do not know anything.' (Kraft 1963: 25)

The copula is also used when there are more than two nouns in a sequence, and where the absence of copula would have resulted in an uninterpretable sequence of three nouns N–N–N:

- (33) *dòrina dabbà ce màì lafiyà*
 hippotamus animal COP.F one.having health
 'Hippotamus is an even-tempered animal.' (Kraft 1963: 67)

The copula is also used in negative clauses:

- (34) *duk dabbà-r dà kè dà wannàn nauyi bà*
 any animal-GEN REL REL.PRES ASSC DEM weight NEG
k'àrama-r dabbà ba cè
 small-GEN animal NEG COP.F
 'Any animal that has such a weight is not a small animal.' (Kraft 1963: 83)

6.9. Conclusions about Chadic

Chadic languages, spoken in a geographically compact area, display different modifying structures and different equational structures. Nevertheless, each correlation of the modifying construction and the equational clause structure confirms our hypothesis, in that if the modifying construction is coded by apposition with no other changes, the equational function requires a marker, which in these languages is a copula. If, on the other hand, the modifying construction uses a marker other than simple apposition, the equational function is coded by apposition of the two NPs. The copula may be used in equational clauses even if the modifying construction is marked by some means other than apposition. In these instances, however, the copula carries a specific function other than coding the equational function; often it codes focus, but it may code other functions such as indicating which NPs in an N–N–N construction are members of the equational clause.

7. Crosslinguistic findings

Our crosslinguistic findings are summarized in Table 2. The table addresses two questions: whether the NP–NP construction has a modifying function or an equational function, and whether a language has a copula in the present tense

Table 2. *Summary of crosslinguistic data*

Language	NP-NP Modifying	NP-NP Predicative	Equational Copula
French	yes*	no	yes
English	yes	no	yes
Hausa	no	yes	no*
Mupun	yes	no	yes
Miya	no	yes	no
Gidar	yes	no	yes
Mina	no	yes	no
Hdi	no	yes	no
Lele	yes	no	yes
East Dangla	no	yes	no
Kanuri	no	yes	no
Lango	no	yes	no
Kisi	no	yes	no
Athpare	no	yes	no
Drehu	yes	no	yes
Mandarin	no	yes	no*
Classical Chinese	no	yes	no*
Japanese	no	yes	no
Mongolian	yes	no	yes
Turkish (spoken)	no	yes	no
Yidiny	no	yes	no*
Gooniyandi	no	yes	no*
Kayardild	no	yes	no
Quechua	yes	yes	no*
Tümpisa	no	yes	no*
Lavukaleve	no	yes	no
Russian	no	yes	no
Polish (literary)	no	no	yes
Polish (non-literary)	no	yes	no*
Kilivila	no	yes	no
Pipil	no	yes	no
Siroi	yes	no	yes
Limbu	yes	no	yes

equational clauses. The column labeled “Equational Copula” indicates whether the copula has the function of coding equational predicates. It does not indicate whether the copula occurs in equational clauses. For the sake of comparison we have also included the data for Chadic languages discussed previously. As in Table 1, “no” means that the construction does not exist, and “yes” means that it does exist. When an asterisk appears, the “yes” or “no” is qualified and is

explained when the details of the language are discussed later in the paper.

To illustrate the interdependency between equational clauses and modifying constructions we have selected languages for which relevant and reliable information is available and for which we can check our data for both supporting and contradicting evidence. The data in the chart represent the results of our interpretations. For the justification of these interpretations see the comments below on individual languages.

7.1. *Kanuri (Nilo-Saharan)*

Kanuri confirms our hypothesis. Equational clauses do not have a copula, and the identity of constituents is coded by juxtaposition, viz. subject–predicate:

- (35) a. *Ali bàrè má*
 Ali farmer
 'Ali is a farmer.'
- b. *shí málàm mòwòntíbè*
 he teacher school
 'He is a school teacher.'
- c. *Músá díwì*
 Musa bad
 'Musa is a bad man.' (Hutchison 1981: 173)

The modification of one noun by another involves the use of postpositions following the modifier:

- (36) *f̄òr Àlì-bè*
 horse Ali-POSTP
 'Ali's horse'

There are a few instances of an NP–NP modifying construction that would seem to provide counterevidence to our hypothesis. However, the apparent counterevidence may well represent derivational constructions where the head can only be the words *kâm* 'person' or *âm* 'people', and these constructions designate men or people having certain attributes: *kâm kàm̀bùwá* 'blind person', *âm kàm̀bùwá* 'blind people', *kâm f̄èrmá* 'a person owning a horse' (Hutchison 1981: 173).

7.2. *Lango (Nilo-Saharan)*

The equational clause in Lango does not have a copula, but the nominal predicate codes the person of the subject and also the aspect:

- (37) *án àdàktàl*
 I 1SG-doctor-HAB
 'I am a doctor.' (Noonan 1992: 144)

There is also a type of equational clauses that could be interpreted as having a copula realized by the form *én*, identical with the 3rd person pronoun. The marker *én* is, however, a means of coding the definiteness of the nominal predicate. “In identification sentences, the third person singular independent pronoun *én* is frequently interposed between the subject and the predicate nominal when the predicate nominal has a definite interpretation” (Noonan 1992: 145):

- (38) a. *mân én gwôkk à dákô òkwàlò*
 this is dog ATTPART woman 3S-steal-PERF
 ‘This is the dog that the woman stole.’
 b. *mân gwôkk*
 this 3SG-dog-HAB
 ‘This is a dog.’ (Noonan 1992: 145–146)

Given the definiteness function of the form *én* it is clear that it does not code the subject-predicate relationship, but rather referentiality.

The lack of a copula in equational clauses predicts, based on our hypothesis, that attributive constructions will use a marker of modification. As this is indeed the case, Lango supports our hypothesis. The attributive constructions in Lango have the order Head–Attributive noun, with the markers of modification inserted in between:

- (39) *gwôkk à lócè*
 dog ATTPART man
 ‘the man’s dog’ (Noonan 1992: 154)

So, both Nilo-Saharan languages, Lango and Kanuri, are similar in that the equational clauses have no copula, and the modifying constructions have a marker of modification. Moreover, the two languages provide further support of our hypothesis because they are two related languages which use different constructions to code NP modification.

7.3. *Kisi (Atlantic, Niger-Congo)*

Kisi has a class system in which modifying constructions are coded through the apposition of the head and modifier, and both the head and the modifier have the same class coding as determined by the class of the head. The head precedes the modifier. All modifiers, including adjectives, possessives, numerals, and demonstratives, code the class of the head noun through a system of suffixes. Moreover, the head also has the class marker:

- (40) a. *mèŋ-mà sàà-áŋ*
 water-CL Saa-CL
 ‘Saa’s water’

- b. *lééŋ-là sàà-láŋ*
 cutlasses-CL Saa-CL
 'Saa's cutlasses'

Equational clauses with adjectival or nominal predicates do not have to include a copula. The copula, however, must be used in focus constructions (Tucker Childs, p.c.). If the copula does occur, the clause has the form Subject–Copula–(Pro)–Noun, where the pronoun (PRO) codes the class of the subject:

- (41) a. *youwo hoo ndu o humbu*
 bird this PRO(EMPH) PRO white
 'This bird, it's white.'
- b. *youwo hoo ndu o humbu o có ni*
 bird this PRO(EMPH) PRO white PRO COP FOC
 'It's that this bird it's white.'
- c. *tàmbà có màsàá*
 Tamba COP chief
 'Tamba is a chief.' (Childs 1995: 327, 253)

The difference between the N–N modifying construction and the equational clause is that in equational clauses there is no suffix after the predicate adjective or noun, while such a suffix is present in modifying constructions. The Kisi data support our hypothesis: the copula is used to code focus in equational clauses, but otherwise is not present in equational clauses, and modifying constructions are coded by a class marker.

7.4. Athpare (Tibeto-Burman)

Athpare confirms our hypothesis since equational clauses with two nouns are marked by apposition, while modifying constructions include marking on the nouns. Ebert (1997: 110) states: "Identificational sentences need no copula, but some speakers use *lena*, which is attested only in this form". Note that when the equational clause does not have a subject, the form *lena* is used in affirmative clauses.

- (42) a. *hitna ka-pit-i? – makhak.*
 that 2POSS-cow-Q is.not
 'Is that your cow? – No.'
- b. *na aya ochorwat capara le-na*
 this my new kodali(hoe) is-NML
 'This is my new kodali.'
- c. *un dhani le-na? – wai-na.*
 he rich is-NML is.not-NML
 'Is he rich? – No.'

- d. *yana kisan makhak, hitna mastor lena.*
 that farmer is.not that teacher is
 'He is not a farmer, he is a teacher.' (Ebert 1997: 110)

Modifying constructions involving two nouns have either a locative or a genitive marker in between the head and the modifier:

- (43) *aniya des-ni-na raja*
 our.EXCL country-LOC-NML king
 'the king of/in our country' (Ebert, p.c.)

Furthermore, most constituents can be attributed with a nominalizer, so that 'village girl' can be rendered either [village-GEN girl] or [village-LOC-NML girl] (Ebert, p.c.).

7.5. *Drehu* (Oceanic, Austronesian)

In *Drehu*, the appositional structure NP–NP is available for the derivation of names of body parts: *pene* 'hair', *mek* 'eye' → *penemek* 'eyelash'; *fene* 'lower part', *im* 'hand' → *feneim* 'palm' (Moyses-Faurie 1983: 114). This structure can also be used for other modifications involving animate nouns:

- (44) *gutuh trahmany*
 chicken man
 'cock' (Moyses-Faurie 1983: 167)

In other modifying constructions, however, nouns are in a dependent form, the structure being Modifier–Head or Head–Modifier with various markers inserted in both orders.

The equational clause has the form: topicalizer *ame* followed by the subject, followed by *tre*, which Moyses-Faurie glosses as 'alors', followed by the predicate noun:

- (45) a. *ame la sinöe celë tre nu*
 TOPZR ART tree this alors coco.tree
 'This tree is a coco tree.'
- b. *ame la ejen la nöj i angaatr tre*
 TOPZR ART name ART country of 3PL alors
Drehu
Lifou
 'The name of their country is Lifou.' (Moyses-Faurie 1983: 197)

Therefore, *Drehu* supports our hypothesis in that the modifying construction and the equational clause occupy different formal niches.

7.6. Chinese (Sinitic)

Classical Chinese and modern Mandarin both support our hypothesis because the modifying construction is coded by grammatical morphemes, whereas equational clauses are coded by juxtaposition. Modifying constructions involve genitive or associative particles.

The possessive is indicated by *zhi* in Classical Chinese (46a) and *de* (46b) in modern Mandarin:⁹

- (46) a. *wang zhi zhu chen*
king GEN various minister
'your Majesty's various ministers' (Meng 1A.7, Pulleyblank 1995: 61)
- b. *wǒ kàn-guo nǐ-de nèi-běn shū*
I read-EXP you-GEN that-CL book
'I have read that book of yours.'

Non-possessive modification in modern Mandarin, termed "associative" in Li & Thompson (1981), is also coded by *de*, which functions here as an associative marker rather than a genitive marker:

- (47) a. *yèwǎn-de tiānkōng*
night-ASSC sky
'the nighttime sky'
- b. *jiàokēshū-de shōujù*
textbook-ASSC receipt
'a receipt for the textbook' (Li & Thompson 1981: 115)

9. Both *zhi* in Classical Chinese and *de* in modern Mandarin may be omitted but the conditions under which those omissions occur are not clear:

- (i) a. *lao yu wang shi*
labor in king business
'They labor in the king's business...' (Meng 5A.4, Pulleyblank 1995: 6)
- b. *wǒ kàn-guo nǐ nèi-běn shū*
I read-EXP you that-CL book
'I have read that book of yours.' (Li & Thompson 1981: 116)

One possible explanation for this phenomenon that Li & Thompson (1981: 169) offer is that the omission of the genitive particle is constrained by whether the NP involved is a topic or a comment. Omission of the genitive particle is acceptable if the NP is a topic, but it becomes ungrammatical if the NP is a comment or a phrase functioning as an object in the sentence. Our attempts to elicit grammaticality judgments with respect to omitted genitive particles yielded mixed results. We conclude that there must be some pragmatic constraints that allow the omission, but we are unable to formulate those constraints.

In Classical Chinese, appositional constructions with a pause between the two NPs are used to topicalize one of the NPs:

- (48) a. *wo, yi-ren*
me one-person
'me, the Lonely One' (Dah Yu Diing)
- b. *nei guang-lie kao Wu wang*
your radiant-deeds father Wu king
'your father of glorious prowess, King Wu'
(Luoh Gaw 50.28, Dobson 1962: 32–33)

The absence of a pause between two juxtaposed NPs in Classical Chinese codes nominal compounding, limited to a specific class of determinate words, those specifying "quasi attributes" (Dobson 1959). Vocational names are compound nouns consisting of a modifying word which refers to a tool or material relevant to a given vocation, followed by a word that means "person" or "workman":

- (49) a. *pi guan*
fur cap
'a fur cap'
- b. *yu ren*
jade man
'a lapidary' (Mencius 1B.9.3, Dobson 1959: 24)

Since the modifying function of apposition is very limited, apposition could be available for coding the equational function. This is indeed the case in modern Mandarin:

- (50) a. *zhāng sān shāndōng rén*
Zhang San Shandong person
'Zhang San is from Shandong.'
- b. *zuótiān yìntiān*
yesterday cloudy.day
'Yesterday was a cloudy day.'
- c. *yúehān hǎo rén*
John good person
'John is a good person.'

The copula can be used in equational clauses in modern Mandarin, but according to Bisang (1998) it indicates a type of focus construction. The copula occurs between the subject and the predicate:

- (51) *wǒ shì bù chī lá de, tā shì shénme dōu kěyǐ chī*
 I be not eat hot GEN 3SG be what all can eat
 [in a restaurant after some confusion as to who can eat what:]
 'It's that I can't eat hot food, and s/he can eat anything.'
 (Li & Thompson 1981: 152–153)

The copula is also used in Mandarin in subjectless equational clauses, as in Gidar and Hdi. In the following example, 'other people' is a topic (the interpretation of the function of the clause and copula is ours):

- (52) *rénjia . shì fēngnián*
 other.people be bumper.year
 'For those people, it was a bumper year.' (Hashimoto 1969: 86)

As in Mandarin, the copula in Classical Chinese codes focus (Dobson 1962, Bisang 1998). Thus Dobson (1962: 93) notes that in Classical Chinese, when a statement "is unqualified or unemphatic, the simple apposition of the two terms serves to mark the relationship between them, but when qualifications (mood), emphasis (in antithetical forms), and so on are introduced, copulas occur between the two terms."

In Classical Chinese, the copula particle *ye* occurs at the end of an equational clause:

- (53) *konzi xián rén ye*
 Confucius virtuous person PRT
 'Confucius [is] an able and virtuous person.' (Li 1993: 278)

If *ye* is not present in an equational clause, a different copula, such as *wei*, may mark the relationship:

- (54) *shi wei tian ming*
 this COP heaven command
 'This is/was the ordinance of Heaven.' (Shujing, Dobson 1962: 95)

The copula is, therefore, not a coding means for the equational clause in either Classical Chinese or modern Mandarin. As in many other languages, the copula is a coding means for focus.

7.7. Japanese

Japanese confirms our hypothesis in that the modifying construction involves a marker of modification, while the equational clause does not obligatorily have a copula.

Standard grammars of Japanese state that the copula is used in equational clauses, but the copula is an optional element of these clauses (Mills 1974: 16–18):

- (55) a. *Sono hito wa Nihonjin da + mitai*
 that person TOP Japanese COP like
 'That person is a Japanese.' + 'like'
- b. *Sono hito wa Nihonjin mitai*
 that person TOP Japanese like
 'That person is like a Japanese.' (Mills 1974: 16–18)

In natural discourse, there are many examples of equational clauses without a copula:

- (56) *Mondai ga mondai [na] daké ni ...*
 problem NOM problem COP just DAT
 'Just because the problem is a problem ...' (Martin 1975: 658)

Furthermore the absence of a copula in equational clauses is recorded in all historical periods of Japanese:

- (57) a. *Yo no naka ha kazu-naki mono ka*
 world GEN inside TOP number-NO thing Q
 'This world is an ephemeral thing, isn't it?'
- b. *Wotomera ha ama-wotome ka-mo*
 maiden.PL TOP heaven-maiden perhaps
 'I wonder if these maidens are the heavenly maidens.'
 (Nara period, Mills 1974: 44)

Just as in the Chadic languages Gidar and Hdi, the copula must be used in subjectless equational clauses with one noun only:

- (58) *Anata, dare?*
 you.SG who
 'Who [are] you?'

The optionality of copula in equational clauses indicates that its function is other than coding predication. The copula in Japanese is a form to which suffixes showing mood, tense, and aspect can be attached. The copula can be omitted if the speaker does not feel it necessary to indicate mood, aspect, or tense (Mills 1974: 19).

Modifying constructions involve the use of the genitive marker *no*:

- (59) *yottu-mózi no hígo, waigo ga*
 four.QUANTIFIER-letter GEN vulgarisms obscenities NOM
 'four-letter vulgarisms and obscenities ...' (Martin 1975: 158)

The genitive marker *no* may be omitted, producing a potential ambiguity between the equational clause and the modifying construction. However, the natural discourse data where there is no genitive marker *no* in the modifying construction clearly rule out the equational function interpretation because of the presence of “compound accentuation”:

- (60) a. *rainen [no] hazime kara ...*
 next.year GEN beginning from
 ‘From the beginning of next year ...’ (Martin 1975: 1048)
- b. *táido [no] ikán ni yotte kimaru*
 attitude GEN how DAT basis decide.PRES
 ‘It is decided on the basis of how one’s attitude is.’ (Martin 1975: 1049)

These modifying constructions without the genitive case marking are not identical with equational clauses because of a difference in prosodic properties, such as pauses and compound accentuation (Martin 1975: 158).

7.8. Mongolian (Altaic)

Mongolian supports our hypothesis: a modifying construction involving two nouns is coded by juxtaposition while the equational clause includes a copula. The equational clause contains a predicate, which is a noun, pronoun, or adjective in subject or direct object form, followed by a copula, which may be omitted (Poppe 1970: 148). The copulas that appear in equational clauses are the finite forms of the verbs *bol-* ‘to become’ and *bae-* ‘to be’, or *yum* ‘is’ or *møn* ‘is indeed’ (in emphatic speech). When the noun functions as a predicate, there is no necessary agreement in case, but there is number agreement:

- (61) a. *enə nom soniŋ baenə*
 this book interesting COP.PRES
 ‘This book is interesting.’ (Poppe 1970: 148)
- b. *enə xerəg xecuuu yum*
 this affair difficult COP
 ‘This affair is difficult.’ (Poppe 1970: 148)

One copula, *møŋ*, has a focusing function, as evidenced by the translation of the following:

- (62) *dorji manae erxəlogči møŋ*
 Dorji our.EXCL manager COP
 ‘Dorji is indeed our manager.’ (Poppe 1970: 148)

Modifying constructions are coded by juxtaposition, with the modifier preceding the head:

- (63) a. *arbəŋ xəxuuur arxi*
 ten wineskin alcohol
 'ten wineskins of liquor' (Poppe 1970: 143–144)
- b. *ulaambaatar xotə*
 Ulan Bator city
 the city of Ulan Bator' (Poppe 1970: 144)
- c. *xuuŋ buɣədə*
 man/person total.number
 'all people' (Poppe 1970: 144)

Possessive constructions involving two nouns must have the genitive case marker *ij* added to the possessor:

- (64) *bagši-ij nom*
 teacher-GEN book
 'the book of the teacher' (Poppe 1970: 144)

7.9. Turkish (Altaic)

Turkish confirms our hypothesis in that the copula in equational clauses is not obligatory, and the modifying constructions are overtly marked either by the "genitive suffix" or "possessive suffix".

The copula *-dır* (with its vowel-harmonic variants *-dir*, *-dür*, *-dür*) is used in writing and formal speech in equational clauses; in ordinary speech, however, the copula is not used in simple 'A=B' clauses (glosses in this section by Z. F., H. K., and A. M.):

- (65) a. *kız-ın ad-ı, Fatma-'dır* (written)
 girl-GEN name-3.POSS Fatima-COP
- b. *kız-ın ad-ı Fatma* (spoken)
 girl-GEN name-3.POSS Fatima
 'The girl's name is Fatima.' (Lewis 1967: 97)
- (66) a. *enerji kaynak-lar-ımız bol-dür* (written)
 power source-PL-1PL.POSS abundant-COP
- b. *enerji kaynak-lar-ımız bol* (spoken)
 power source-PL-1PL.POSS abundant
 'Our sources of power are abundant.' (Lewis 1967: 97)

According to Underhill (1976: 208), the morpheme *-dır* emphasizes the truth of the statement being made. Usually it is used with the 3rd person; it may be used with the narrative past affix to "remove the connotations of doubtfulness" from this tense.

The copula is obligatory in subjectless equational clauses:

- (67) a. *yaman bir adam-dir*
 remarkable one man-COP
 'He is a remarkable man.' (Lewis 1967: 97)
- b. *yaman bir adam-dir, amca-nız*
 remarkable one man-COP uncle-2PL.POSS
 'He is a remarkable man, your uncle.' (Lewis 1967: 97)

The copula is also obligatory when there is a possible confusion between an equational clause and an appositional construction:

- (68) *en çok sevdiğim şair Nedim-dir*
 most favorite.1SG.POSS poet Nedim-COP
 'My favorite poet is Nedim.'
 (topicalization: 'My favorite poet, Nedim ...')

In colloquial speech, there is the alternative of using the 3rd person pronoun *o* instead of *-dir*: *o, yaman bir adam* or *yaman bir adam, o* 'he is a remarkable man', compare (67a).

Aside from subjectless equational clauses, the use of *-dir* in informal speech indicates either emphasis or an obligatory conclusion:

- (69) *vesika kasa-da-dir*
 document safe-LOC-COP
 'The document is in the safe.' (written)
 'The document must be in the safe.' or (spoken)
 'The document IS in the safe.' (Lewis 1967: 97)

Possession is marked by two suffixes: the genitive case suffix *-in* attached to the possessor, and the 3rd person possessive suffix *-i* attached to the possessum:

- (70) *Mehmed-in el-i*
 Mehmed-GEN hand-3.POSS
 'Mehmet's hand' (Underhill 1976: 91)

Our interpretation of Turkish data as given in Underhill (1976) is that the non-possessive modification of one noun by another can be coded in two ways. If the two nouns have different semantic scopes, the noun with the broader scope (the modifier) is first and the noun with the narrower scope, the head, is second. Our working definition for the terms "broad" and "narrow" scope hinges on the number of potential referents. Thus a common noun has a broader scope than a proper noun as seen in the following example:

- (71) *dost-um Ahmet*
 friend-1SG.POSS Ahmet
 'my friend Ahmet' (Underhill 1976: 209)

The equational clause has the noun with the narrower scope as the subject, and the noun with the broader scope as the predicate:

- (72) *Ahmet dost-um(-dur)*
 Ahmet friend-1SG.POSS(-COP)
 'Ahmet is my friend' (Underhill 1976: 209)

If the two nouns have inherently equal scope, then the head will be marked by the morpheme *bir* 'one' functioning as an indefinite article. The presence of this morpheme has therefore the function of narrowing the scope of the head:

- (73) a. *taş-tan bir ev*
 stone-ABL one house
 'a stone house' (Underhill 1976: 209)
 b. *ev taş-tan*
 house stone-ABL
 'The house is of stone.' (Underhill 1976: 209)

7.10. *Yidiny* (Australian)

There are several types of constructions in *Yidiny* that involve two nouns in apposition. In one construction the first noun is a classifier, and the choice of classifier nouns is limited to about twenty generic nouns (Dixon 1977: 483):

- (74) *mija gangu:l*
 animal-ABS wallaby-ABS
 'the animal wallaby' (Dixon 1977: 480)

The second type is a part-whole modifying construction, with the modifier preceding the head, where the whole (modifier) precedes the part (head):

- (75) *pundu | bama dungu numaj*
 hey! person.ABS head.ABS move.about.PRES
 'Hey! That's a person's head moving about.' (Dixon 1977: 248)

The function of these two types of constructions is never ambiguous because it is fully determined by the highly constrained lexical features of the modifiers and their relationship to the lexical features of the heads. The interpretation of these constructions as equational clauses would result in a nonsensical reading. Thus for the classification construction, the reading would have a subject with much broader scope than the predicate, such as *'animal is wallaby'. For part-whole constructions, the equational reading will again produce a nonsensical interpretation where part and whole are predicated as being equal, such as 'person is head' or 'head is person'.

For modifying constructions, where the nouns are not lexically constrained, the modifying construction is coded by the use of the genitive suffix *-ni* ~ *-nu*:

- (76) *yijū gala guran ŋuŋdu:n bama:n*
 this.ABS spear.ABS long.ABS that.GEN.ABS person.GEN.ABS
mayaɾani
 initiated.man.GEN.ABS
 'This long spear belongs to that Maya:ɾ man.' (Dixon 1977: 358)

Given the limited nature of modifying constructions involving two nouns, and the productive means of using the genitive case marker for other modifying constructions, our hypothesis predicts that the NP–NP niche is available for the equational function, and that there need not be a copula. And this is indeed the case in Yidiny, as in most Australian languages.

In equational clauses, the subject noun is in the absolutive case, a subject pronoun is in the nominative case, and the predicate, which may be a noun or adjective, is presumably also in the absolutive case (there are no glosses in the original): *ɖugi yijū ɖundu* 'this tree is [just] a stump', *mayi mamba* 'this fruit is sour' (Dixon 1977: 271).

An equational clause is used in reporting the proper names of objects, such as when a hero in a story names a place based on a geographical feature or something that occurred at the place: *bulmba yijū balbu:ɾu | balbun ɖana:n* 'this place [was called] Balbu:ɾu: the Balbun tree grew [there]' (Dixon 1977: 272).

7.11. *Gooniyandi (Australian)*

In Gooniyandi, equational clauses do not have a copula, and the predicate follows the subject:

- (77) *ngaddagi thangarndi Gooniyandi*
 my language Gooniyandi
 'My language is Gooniyandi.' (McGregor 1990: 295)

According to our hypothesis, modifying constructions, including possessive constructions, should have a marker. This is indeed the case. McGregor (1990: 252) states that "[i]n NPs expressing both alienable and inalienable possession, the possession is referred to either by an embedded dative PP [...], or by an oblique case-form of the free pronominal":

- (78) a. *ngaddagi marla / thadda*
 my hand / dog
 'my hand/dog' (McGregor 1990: 252)
 b. *ngooddoo-yoo goornboo marla / thadda*
 that-DAT woman hand / dog
 'that woman's hand/dog' (McGregor 1990: 252)

The only examples that may provide counterevidence for our hypothesis are those where the two nouns are in apposition and do not involve an equational clause:

- (79) *thiddoo nyawa*
 kangaroo tail
 'kangaroo tail' (McGregor 1990: 253)

But, as McGregor explains it, the first noun in such constructions is a classifier. The situation resembles the one described earlier for Yidiny.

7.12. *Kayardild (Tangkic subgroup, Australian)*

In Kayardild, neither the attributive nor the predicative constructions have overt morphological markers of their functions. Instead, the functions are coded by the scope of the two NPs involved. If the first NP has a narrower scope than the second NP, the construction is an equational clause. The narrower scope could be determined by the use of determiners:

- (80) a. *dathin-a yakuri-ya katharr-waan-d*
 that-NOM fish-NOM swamp-ORIG-NOM
 'That fish is a swamp fish.'
 b. *dathin-a nguriw dun-marri*
 that-NOM teenage.girl.NOM spouse-PRIV(NOM)
 'That girl is unmarried/has no husband.' (Evans 1995: 315)

If the first NP has wider scope than the second, the construction codes modification:

- (81) *dun-marri nguriwa*
 spouse-PRIV(NOM) teenage.girl
 'unmarried girl'

The data from Kayardild thus support our hypothesis; although there are no morphological markers of type involved, functional coding is achieved through the configuration of NPs based on their scopes. The data from Kayardild may also point to how copulas developed from determiners. The marker of the narrower scope may become a copula through the reanalysis of its function.

7.13. *Quechua (Quechuan)*

If one were to follow the descriptions of Quechua in Weber (1989), Parker (1969), Grondín (1990), the statements in those descriptions contradict our hypothesis. Therefore we decided to examine the relevant information on Quechua more closely, using natural language data whenever possible. There is considerable variation among dialects of Quechua with respect to attributive

and equational constructions. In some dialects our hypothesis is supported by the data, but in one dialect it is contradicted. We include here data from the following dialects: Ayacucho, Cuzco, Potosí, Cochabamba, and Santiago del Estero. In all the dialects of Quechua, there are modifying constructions coded by juxtaposition, i.e., of the type NP–NP, where the first NP is a modifier. The following sentences from three different dialects illustrate this point (the relevant sequences are in roman):

- (82) a. *rima-pu-q-ni-n-taq* *qapari-n chay qhari*
 say-BEN-AG-FILL-3SG-CONJ yell-3SG that man
wawa-ta: Yau mana yacha-q *uchu puna*
 child-ACC INJ not know-AG little mountain
indio-cha
 Indian-DIM
 ‘When he said that he yelled at the boy, ‘You ignorant little mountain boy.’ (Ayacucho; Gibbons & Gibbons 1998a)
- b. *Qosqo llaqta-manta-pacha napayku-ykichis ancha misk’i*
 Cuzco city-ABL-itself greet-1S-2PL.OBJ very sweet
simi-y-wan *llapan teqsimuyu runakuna-ta*
 word-POSS.1SG-INSTR all earth people-ACC
 ‘From the city of Cuzco I greet all the people of the world with my sweet words.’
 (Cuzco; <http://www.unsaac.edu.pe/CUSCO/RunaSimi/>)
- c. *Haku p’isqu runtu mask’a-sun*
 come bird egg look.for-1PL.HORT
 ‘Come, let’s look for bird eggs.’ (Bolivia, Potosí; Crapo & Aitken 1986: 44)

This modification differs from possessive constructions in that the possessor is usually marked with genitive case while the possessum has a possessive suffix that agrees in person with the possessor.

- (83) *Ñawpaq ka-q wasimasi asnu-n-mi*
 first be-AG neighbor donkey-POSS.3SG-EV
yacha-ku-rqa iskay ka-q wasimasi-q sara
 know-REF-PAST second be-AG neighbor-GEN corn
tarpu-y-ni-n-man *yayku-y-ta*
 plant-INF-FILL-POSS.3SG-DIR enter-INF-ACC
 ‘A man’s donkey was always getting loose, and going into the neighbor’s corn field.’ (Ayacucho; Gibbons & Gibbons 1998a)

There is considerable variation with respect to the structure of equational clauses. Some equational clauses appear to have a structure identical to that of

modifying constructions, thus contradicting our hypothesis; other types have a marker, albeit not necessarily a copula, thus confirming the hypothesis. Many varieties of Peruvian Quechua obligatorily include evidential markers in all main clauses, except in the imperative mood. These evidential suffixes are *-mi* (with various allomorphs *-m*, *-n*, *-min*) for first hand experience and *-si* (also with various allomorphs *-s*, *-sina*) for hearsay. These evidential suffixes can act as markers of predication in equational clauses:

- (84) a. *Warmi-y-pas kamachiq-ni-yki-m.*
 wife-POSS.1SG-also boss-FILL-POSS.2SG-EV
 'My wife is also your boss.' (Ayacucho; Gibbons & Gibbons 1998b)
- b. *San Antuñu-n suti-n.*
 San Antuño-EV name-3SG
 'Its name is San Antuño.' (Cuzco,
<http://www3.rcp.net.pe/QUECHUA/CAPITULO4/8.htm>)
- c. *Warmi-kuna-m aswan ashkha qhawaqñawsa*
 woman-PL-EV more a.lot illiterate
 'The women are a lot more illiterate.' (Cuzco; *Radio Tinku*,
 March 9, 1998)
- d. *Qhari-kuna-lla-m alcalde.*
 man-PL-DEL-EV mayor
 'Only men are mayors.' (Cuzco; *Radio Tinku*, March 9, 1998)

In some dialects, the subject of the equational clause is marked by a topicalization marker, a marker that is absent in the modifying construction:

- (85) a. *Pay-qa don Alberto.*
 he-TOP don Alberto
 'He is don Alberto.' (Cochabamba; Grondín 1990: 32)
- b. *Kay-qa aqha.*
 this-TOP corn.beer
 'This is corn beer.' (Cuzco; Juan Antonio Manya, p.c.)

In the Santiago del Estero dialect spoken in Argentina, the copula is obligatory in equational clauses in all persons:

- (86) *Erramyenta-s-ni-y-qa ka-n lima-s tukuy laya*
 tool-PL-FILL-POSS.1SG-TOP be-3SG file-PL all kinds
tamaño-s
 size-PL
 'My tools are files, of all sizes.' (Santiago del Estero; Cerrón-Palomino 1987: 395)

However, in some dialects of Quechua, such as the Bolivian and Ecuadorian varieties (Muysken 1977: 28), the evidential suffixes are not obligatory, and the equational clause has the same structure as the modifying construction:

- (87) *Kay runtu-kuna chiwanku runtu.*
 this egg-PL blackbird egg
 'These eggs are blackbird eggs.' (Bolivia, Potosí; Crapo & Aitken 1986: 44)

Thus, the Bolivian dialects contradict our hypothesis in that the NP–NP construction can be used in both modification and equational clauses.

In at least one dialect where equational clauses have no copula, existential clauses do have a copula. This fact indicates that there are three functional domains which compete for the same formal niche:

- (88) *Kay p'uyñu aqa-lla-m ka-n.*
 this urn corn.beer-DEL-EV exist-3SG
 'There is just this pitcher of corn beer.' (Ayacucho; Gibbons & Gibbons 1998a)

In conclusion, Quechua contains NP–NP modification which is coded by juxtaposition. The Santiago del Estero dialect requires a copula in equational clauses. Many Peruvian varieties of Quechua also confirm our hypothesis since the obligatory evidential suffixes can serve as markers of predication for equational clauses. In certain varieties of Quechua (Bolivian) the evidential markers are not obligatory. These dialects contradict the hypothesis put forth in this paper. At present we have no information regarding the prosodic structures of equational and modifying constructions nor an explanation regarding the evolution of these constructions in Quechua dialects.

7.14. *Tümpisa (Panamint) Shoshone (Uto-Aztecan)*

Data from Tümpisa Shoshone support our hypothesis. The equational clause does not obligatorily contain a copula, and the modifying construction is coded by the agreement system between the head and the modifier. The copula *naa* 'be' is used in both linking and existential sentences, and it may be used in all tenses and aspects; however, it is usually omitted in the simple present or narrative present unless the speaker wants to indicate an aspectual distinction using an aspectual suffix on the verb:

- (89) a. *Satü wihnu niam petü*
 that then my daughter
 'That's my daughter then.' (Dayley 1989: 28)

- b. *Nümmi appü utü*
 our.EXCL father that
 'That's our father.' (Dayley 1989: 28)

To code tense and aspect, the copula is necessary and may not be omitted:

- (90) *Tammi appü naammaa satü wihnu*
 our.INCL father was that then
 'That was our father then.' (Dayley 1989: 28)

In modifying constructions, the head and the modifier must agree with respect to case marking:

- (91) *Nü püyünna pühi punikkappühantü*
 I duck-OBJ pelt-OBJ saw
 'I saw a duck pelt.' (Dayley 1989: 275)

7.15. *Lavukaleve (isolate, Solomon Islands)*

In Lavukaleve, modifying constructions involving two nouns are coded with the genitive, while equational clauses involve juxtaposition of the two NPs.

According to Angela Terrill (personal communication), "usually the only possible way to have two nouns within one NP is to put them in a head-modifier relationship using a possessive construction". For example (all data: Angela Terrill, p.c.):

- (92) a. *Aka foiga molio la o*
 then PRO.MED.SG.N tree.species(F) SG.F.ART and
beko mina an'kav o-bek na
 stone(F) um pumice(M) 3SG.POSS-stone(M) SG.M.ART
nego vo-e-me-n hide fi
 float 3PL.O-SBD-HAB-ADMON thus 3SG.N.FOC
 'So [they would take] the molio and the stone – what? –, the pumice stone that can float, ...' (cs2 024)
- b. *Maiva | mate o-ta*
 I.don't.know war(N) 3SG.POSS-time(M)
a-e-sia-ge ve
 3SG.M.O-SBD-do-ANT or
 'Maybe because it was war time.' (co2 072)
- c. *Mola o-fo'foira o-lo-e-n*
 canoe(n) 3SG.POSS-work(F) 3SG.POSS-finish-NML-LOC
me-a heo
 SPEC-SG.F 3SG.F.EFOC
 'That's the end of the canoe work.' (cp 040)

Although it is possible to have two nouns juxtaposed while not in a possessive construction, these phrases are always lexicalized phrases:

- (93) a. *mo'sil savu*
 shore edge
 'sea shore'
 b. *ta'rai tail*
 prayer house
 'church'

There is no copula in Lavukaleve, so equational clauses simply involve two NPs, as well as any adjuncts. Focus marking is also often present in equational clauses:

- (94) a. *Lafi ga laketei fi*
 water(N) SG.N.ART life(N) 3SG.N.FOC
 'Water is life.' (e1 016b)
 b. *Todou na fin foina*
 Todou(M) SG.M.ART 3SG.M.FOC PRO.MED.SG.M
marai hin
 warrior(M) 3SG.M.EFOC
 'Todou was a warrior.' (e1 052d)
 c. *O! Fova suni nga-tulav*
 oh PRO.PROX.PL all 1SG.POSS-children(PL)
hova!
 MOD.PROX.PL
 'Oh! These are all my children!' (jn2 038)
 d. *Foiga ma-tail hi*
 PRO.MED.SG.N 3PL.POSS-house(N) 3SG.N.EFOC
 'That's their house.' (fk 025)

8. Absence of the NP–NP formal niche

Now we would like to examine languages that are usually said to have a copula in equational clauses and some other marker in modifying constructions, i.e., languages that would appear not to make use of the formal niche NP–NP. We limit our investigation to two well-described Indo-European languages, Polish and French.

8.1. Equational clauses in Polish

Polish is similar to Russian in that the modifying construction uses either an adjectival form of the modifier or the genitive case marker. Yet, unlike in Russian, the equational clause in Polish does have a copula in the present tense, at

least in literary Polish. One of the environments that does not allow a copula in the literary variety of Polish is in comments on a topic. Such clauses cannot have a copula:

- (95) *nie było mnie, to prawda*
 NEG be 1SG.ACC DEM truth
 'I wasn't there, it is true.' (Sources)

In the past tense, the copula must always be used, because this is the means to code the tense:

- (96) *dżiki naród był*
 wild nation be.PAST.3M.SG
 '[They] were wild people.'

In non-literary varieties of Polish, more specifically in the speech of illiterate speakers, the copula is not always used. The following are a few examples from different dialects of non-literary varieties of Polish with adjectival predicates:

- (97) a. *ké to pšéce né ĵego žec*
 but DEM indeed NEG his thing
 'But that is none of his business.' (Nitsch 1960: 85; all glosses and translations for this source by Z. F.)
 b. *to pěkto okropné duže*
 DEM hell awfully big
 'That hell is awfully big.'

Here are examples with nominal predicates without a copula. Note that the subject in such predicates is the neutral singular demonstrative *to*:

- (98) *bo to ĵu--šf'ěty pėter ĵi pavel, a tu dės*
 because DEM already-Saint Peter and Paul and here rain
 'because it is already Saint Peter and Saint Paul[']s day], and here it rains' (Nitsch 1960: 154)

The copula is used, however, when the subject is separated from the predicate by some material that does not belong to either:

- (99) a. *dyž my teš sům řvāla Panu B^uogu*
 because we also be.3PL glory Lord.DAT God.DAT
zdr^uoři
 healthy
 'because we are, thanks to God, healthy' (Nitsch 1960: 35)
 b. *u nas ty hařupska sům řšeckı ĵednāki*
 at 1PL dem.PL houses be.3PL all identical
 'In our villages all houses are identical.' (Nitsch 1960: 37)

Compare equational clauses without a copula from the same speaker:

- (100) *ĭ Costkovĭan caty ŝčyĭŝl'ivy*
 and Costkovian all happy
 'and Costkovian is all happy' (Nitsch 1960: 36)

Here are examples with nominal predicates:

- (101) *a baby tam p"o čni ěe poznať,*
 and women there PREP darkness NEG recognize.3PL.F
že to ěe ěeĭĭχ χtōpi
 COMP DEM NEG their men
 'And the women there did not recognize in the darkness that those are not their men.' (Nitsch 1960: 37)

8.2. Equational clauses in French

Standard grammars of literary French show a picture in which the modifying construction, in this case a genitive, has a particle *de* and the equational clause has a copula, as in the following examples:

- (102) *les tragédies de Racine*
 DEF.PL tragedy.PL PREP Racine
 'Racine's tragedies' (Grevisse 1991: 563)
- (103) *la crainte du Seigneur est le commencement de*
 DEF fear PREP Lord COP DEF beginning PREP
la sagesse
 DEF wisdom
 'The fear of the Lord is the beginning of wisdom.' (Grevisse 1991: 526)

It would appear that the structure NP–NP is not exploited for either the modifying construction or the equational clause. Nevertheless, one can find exceptional, non-productive uses of the NP–NP structure for both equational clauses and the modifying function in newspaper headlines and other instances of telegraphic use of the language. The history of how this particular situation developed in French provides support for the proposed hypothesis about the constraint that the same formal niche cannot be used for different functions.

Latin had several types of equational clauses without copula:

- (104) a. *homo homini lupus*
 man.NOM man.DAT wolf.NOM
 'Man is wolf to another man.'

- b. *mors ultima ratio*
 death ultimate reason
 'Death is the ultimate reason [of everything].' (Both examples
 quoted in Feuillet 1998: 681)

Ernout & Thomas (1972) explain that the omission of the copula in adjectival predicates is quite rare because it would lead to ambiguity between predicative and attributive expressions, whereby *pulchra domus* could be interpreted as either 'the house is beautiful' or 'the beautiful house'. Hence the use of the copula in predicative construction allows this configuration to be used for the modifying function. With respect to nominal predicates, Ernout & Thomas state that the absence of the copula is not a result of ellipsis. When they say that such structures represent 'old expressions', we interpret it to mean that these expressions never had a copula: *amor omnibus idem* 'love is the same for all' (Ernout & Thomas 1972: 146).

In medieval French, possessive constructions, especially those involving human possessors, consisted of apposition of possessum and possessor:

- (105) a. *le cor Rollant*
 DEF heart Roland
 'Roland's heart'
 b. *le cheval le roi*
 DEF horse DEF king
 'the king's horse'
 c. *les quatre fils Aymon*
 DEF.PL four son.PL Aymon
 'the four sons of Aymon' (Grevisse 1991: 573)

Since the niche NP–NP was occupied by one function it could not have been occupied by another. The equational clause therefore had to have some means of coding other than juxtaposition.

Starting in the 14th century, the appositional modifying construction begins to disappear. Nevertheless contemporary French has many instances of fixed expressions containing NP–NP apposition, as well as plenty of evidence of productive use of the NP–NP construction for the modifying function. Here are some examples of fixed expressions:

- (106) a. *la tour Eiffel*
 DEF tower Eiffel
 'Eiffel tower'
 b. *la rue Notre-Dame*
 DEF street Notre-Dame
 'Notre-Dame street' (Grevisse 1991: 573)

Below are examples (Grevisse 1991: 574) of productive use of the NP–NP construction for a modifying function, which Grevisse attributes either to foreign influences or to the ease of use:

- (107) a. *le match France–Belgique*
DEF match France–Belgium
'the France–Belgium match'
- b. *un lavage minute*
INDEF wash minute
'a minute wash [for cars]'
- c. *une poche poitrine*
INDEF pocket breast
'a breast pocket'

Given the fact that the NP–NP structure is not productive for equational clauses in French but is quite productive for modifying constructions involving two nouns, the data in French support our hypothesis.

9. Conclusions

The purpose of this paper was to examine whether the use of some formal niche A for a function X precludes the use of this niche for another function. The definition of a formal niche includes the lexical categories which may participate in a construction, configuration of the lexical categories, and all other coding means that are used, such as prosodic and other phonological means, inflectional means, and syntactic means such as use of adpositions. We have chosen two functions whose realization must minimally involve use of the same lexical categories: modifying constructions involving two nouns, one serving as the head and the other as a modifier, and equational clauses, with one noun serving as a subject and the other as a predicate.

Our findings support the following hypothesis: If the formal niche X is used for function A, it cannot be used for function B. If the equational clause in the language is coded by juxtaposition alone, the modifying function must be coded by some other means. Furthermore, if a modifying function is coded by juxtaposition alone, the equational clause has a copula or some other means of coding the equational function. For this very narrow range of functions, we were able thus to provide a functional explanation for the question of why some languages have a given structure with a given function and other languages do not.

Received: 19 March 2001

Revised: 19 March 2002

University of Colorado

Correspondence address: (Frajzyngier) Dept. of Linguistics, Box 295, University of Colorado, Boulder, CO 80309, U.S.A.; e-mail: zyg.munt.frajzyngier@colorado.edu

Acknowledgments: The work of Zygmunt Frajzyngier on Mina, Gidar, Hdi, and Lele was supported by grants from the National Endowment for the Humanities and from the National Science Foundation. The data on East Dangla come from Erin Shay whose work was supported by the National Science Foundation through a grant to Zygmunt Frajzyngier. We are grateful to Bob Dixon for comments on Yidiny, to Eunil Kim for data on Korean, to Tuul for Mongolian data, to Saeko Ogiwara and Hayashi Makoto for the help with the Japanese data, to Karen Ebert for help in the interpretation of Athpare material, to Claire Moysse-Faurie for the help in the interpretation of Drehu data, and to Angela Terrill for the data and analyses of Lavukaleve. Peter Cole's comments on facts in Quechua prompted us to examine more carefully data from various dialects of Quechua. A preliminary version of this paper was presented at the 3rd International Conference of the Association for Linguistic Typology in Amsterdam, August 1999. We are most grateful to two anonymous readers of this paper for their many helpful comments. As a result of these comments we have added some new analyses and formulations. We alone are responsible for interpretations of the data provided by our colleagues and for the resulting conclusions.

Abbreviations: ABL ablative, ABS absolute, ACC accusative, ADMON admonitive, AG agentivizer (derivational, nominalizer), ANT anterior, ART article, ASSC associative, ATTPART attributive particle, BEN benefactive, CL class marker, COM comment marker, COMP complementizer, CONJ conjunction, COP copula, DAT dative, DEF definite, DEL delimitative, DEM demonstrative, DIM diminutive, DIR directional, EFOC focus marker from *heo* paradigm in Lavukaleve, EMPH emphatic, EV evidential, EXCL exclusive, EXP experiential, EXT extended, F feminine, FILL filler (filler suffix to prevent disallowed consonant clusters), FOC focus marker from *feo* paradigm, GEN genitive, HAB habitual, HORT hortative, HYP hypothetical, IMP imperative, INCL inclusive, INDEF indefinite, INF infinitive (nominalizer suffix), INJ interjection, INST instrumental, INT intention, INTENS intensive, LOC locative, M masculine, MED medial (demonstrative category), MOD modal, N neuter, NEG negative (verb suffix), NML nominalizer, O object, OBJ object, OBL oblique, ORIG origin, PAST past tense marker, PERF perfective, PL plural, POSS possessive, POSTP postposition, POT potential (verb suffix), PREP preposition, PRES present tense, PRIV privative, PRO pronoun, PROX proximal (demonstrative category), PRT particle, Q question marker, REFL reflexive, REFR referential, REL relative, S subject, SBD subordinate (verb prefix), SG singular, SPEC specifier (adjective), TOP topic, TOPZR topicalizer.

References

- Admoni, V. G. (1983). Nulevaja svjazka, svjazočnyj glagol i grammatika zavisimostej. [The zero copula, auxiliary verbs, and grammar of dependencies.] *Voprosy jazykoznanija* 1983–5: 34–42.
- Bisang, Walter (1998). Grammaticalization and language contact, constructions and positions. In Anna Giacalone Ramat & Paul J. Hopper (eds.), *The Limits of Grammaticalization*, 13–58. Amsterdam: Benjamins.
- Cerrón-Palomino, Rodolfo (1987). *Linguística Quechua*. Cuzco, Peru: Centro de Estudios Rurales Andinos "Bartolome de Las Casas".
- Childs, George Tucker (1995). *A Grammar of Kisi: A Southern Atlantic Language*. Berlin: Mouton de Gruyter.
- Cohen, David (1984). *La phrase nominale et l'évolution du système verbal en sémitique*. Leuven: Peeters.
- Cohen, Marcel (1924). *Le système verbal sémitique et l'expression du temps*. Paris: Leroux.
- (1970) [1937]. *Traité de langue amharique*. Paris: Institut d'Ethnologie.
- Crapo, Richley H. & Percy Aitken (1986). *Bolivian Quechua Reader and Grammar-Dictionary*. Ann Arbor, Mich.: Karoma.

- Dayley, Jon P. (1989). *Tümpisa (Panamint) Shoshone Grammar*. Berkeley: University of California Press.
- Déprez, Viviane & Marie-Thérèse Vinet (1992). Une structure predicative sans copule. *Revue québécoise de linguistique* 22: 11–44.
- Dixon, R. M. W. (1977). *A Grammar of Yidiny*. Cambridge: Cambridge University Press.
- Dobson, W. A. C. H. (1959). *Late Archaic Chinese*. Toronto: Toronto University Press.
- (1962). *Early Archaic Chinese*. Toronto: Toronto University Press.
- Ebert, Karen (1997). *A Grammar of Athpare*. München: Lincom Europa.
- Ernout, Alfred & François Thomas (1972). *Syntaxe latine*. Paris: Klincksieck.
- Evans, Nicholas D. (1995). *A Grammar of Kayardild*. Berlin: Mouton de Gruyter.
- Feuillet, Jack (1998). Typologie de 'être' et phrases essives. In Jack Feuillet (ed.), *Actance et valence dans les langues de l'Europe*, 663–752. Berlin: Mouton de Gruyter.
- Frajzyngier, Zygmunt (1993). *A Grammar of Mupun*. Berlin: Reimer.
- (2001). *A Grammar of Lele*. (Stanford Monographs in African Linguistics, 5.) Stanford: Center for the Study of Language and Information.
- Frajzyngier, Zygmunt, Eric Johnston, & Adrian Edwards (2002). A grammar of Mina. Unpublished manuscript.
- Garrigues-Cresswell, Martine & Christophe Weibegué (1981). *Livre de lecture lélé*. Sarh, Chad: Centre d'études linguistiques.
- Gibbons, Ada & Russ Gibbons (1998a). Jokes in Quechua. <http://www.andes.org/jokes.html>
- (1998b). Basic Quechua lessons. http://www.andes.org/q_phrase.html
- Godfrey, John J., Edward C. Holliman, & Jane McDaniel (1992). Switchboard: Telephone speech corpus for research and development. In *Proceedings of ICASSP-92: 17th International Conference on Acoustics, Speech and Signal Processing*, Volume 1, 517–520. Piscataway, N.J.: IEEE Service Center.
- Grevisse, Maurice (1991). *Le bon usage: Grammaire française*. 12th edition (edited by André Goosse). Paris: Duculot.
- Grondín N., Marcelo (1990). *Método de Quechua, Runa Simi*. 3rd edition. La Paz: "Los Amigos del Libro".
- Hashimoto, Anne Yue (1969). The Verb 'to be' in Modern Chinese. In John W. M. Verhaar (ed.), *The Verb 'Be' and its Synonyms: Philosophical and Grammatical Studies*, Volume 4: *Twi, Modern Chinese, Arabic*, 72–111. Dordrecht: Reidel.
- Heine, Bernd, Tom Güldemann, Christa Kilian-Hatz, Donald A. Lessau, Heinz Roberg, Mathias Schladt, & Thomas Stolz (1993). *Conceptual Shift: A Lexicon of Grammaticalization Processes in African Languages*. (Afrikanistische Arbeitspapiere, 34/35.) Köln: Institut für Afrikanistik.
- Hengeveld, Kees (1990). A functional analysis of copular constructions in Mandarin Chinese. *Studies in Language* 14: 291–323.
- Hutchison, John P. (1981). *A Reference Grammar of the Kanuri Language*. University of Wisconsin–Madison, Boston University African Studies Center.
- Jokwani, Mbulelo W. (1997). Identifying, descriptive and associative copulatives in Xhosa: structure and function. *South African Journal of African Languages* 17: 110–113.
- Jungrathmayr, Herrmann (1978). Les langues tchadiques et le proto-tchadique: documentation, analyse et problèmes. In Jean-Pierre Caprile & Herrmann Jungrathmayr (eds.), *Préalables à la reconstruction du proto-tchadique*, 17–30. Paris: SELAF.
- Katz, Aya (1996). Cyclical grammaticalization and the cognitive link between pronoun and copula. Doctoral dissertation, Rice University.
- Kraft, Charles H. (1963). *A Study of Hausa Syntax*. Hartford, Conn.: Hartford Seminary Foundation.
- Kurcz, Ida, Andrzej Lewicki, Jadwiga Sambor, Krzysztof Szafran, & Jerzy Woronczak (1990). *Słownik frekwencyjny polszczyzny współczesnej*. [Frequency Dictionary of Contemporary Polish.] Kraków: Polska Akademia Nauk.

- Lefebvre, Claire & Pieter Muysken (1988). *Mixed Categories: Nominalizations in Quechua*. Dordrecht: Kluwer.
- Lewis, G. L. (1967). *Turkish Grammar*. Oxford: Oxford University Press.
- Li, Charles N. & Sandra A. Thompson (1977). A mechanism for the development of copula morphemes. In Charles N. Li (ed.), *Mechanisms of Syntactic Change*, 414–444. Austin: University of Texas Press.
- (1981). *Mandarin Chinese: A Functional Reference Grammar*. Berkeley: California University Press.
- Li, Feng (1993). The copula in Classical Chinese declarative sentences. *Journal of Chinese Linguistics* 21: 277–312.
- Lichtenberk, Frank (1983). *A Grammar of Manam*. Honolulu: University of Hawaii Press.
- MacDonald, Lorna (1990). *A Grammar of Tauya*. Berlin: Mouton de Gruyter.
- Martin, Samuel E. (1975). *A Reference Grammar of Japanese*. London: Yale University Press.
- Matthews, Peter (1997). *The Concise Oxford Dictionary of Linguistics*. Oxford: Oxford University Press.
- McGregor, William (1990). *A Functional Grammar of Gooniyandi*. Amsterdam: Benjamins.
- McWhorter, John (1994). From focus marker to copula in Swahili. *Berkeley Linguistics Society* 20S: 57–66.
- Mills, David O. (1974). A historical linguistic study of the copula in the Japanese language. Doctoral dissertation, University of Michigan.
- Moyse-Faurie, Claire (1983). *Le drehu: Langue de Lifou (Iles Loyauté)*. Paris: SELAF.
- Muysken, Pieter (1977). *Syntactic Developments in the Verb Phrase of Ecuadorian Quechua*. Dordrecht: Foris.
- Newman, Paul (1977). Chadic classification and reconstructions. *Afroasiatic Linguistics* 5-1: 1–42.
- (2000). *The Hausa Language: An Encyclopedic Reference Grammar*. New Haven: Yale University Press.
- Nitsch, Kazimierz (1960). *Wybór polskich tekstów gwarowych*. [Selection of Polish Dialect Texts.] Warszawa: Państwowe Wydawnictwo Naukowe.
- Noonan, Michael (1992). *A Grammar of Lango*. Berlin: Mouton de Gruyter.
- Parker, Gary J. (1969). *Ayacucho Quechua Grammar and Dictionary*. The Hague: Mouton.
- Pisarkowa, Krystyna (1984). *Składnia historyczna języka polskiego*. [Historical Syntax of Polish.] Wrocław: Zakład Narodowy im. Ossolińskich.
- Poppe, Nicholas (1970). *Mongolian: Language Handbook*. Washington: Center for Applied Linguistics.
- Pulleyblank, Edwin G. (1995). *Outline of Classical Chinese Grammar*. Vancouver: UBC Press.
- Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech, & Jan Svartvik (1983). *A Comprehensive Grammar of the English Language*. New York: Harcourt, Brace, Jovanovich.
- Schuh, Russell G. (1998). *A Grammar of Miya*. Berkeley: University of California Press.
- Senft, Gunter (1986). *Kilivila: The Language of the Trobriand Islanders*. Berlin: Mouton de Gruyter.
- Stassen, Leon (1994). Typology versus mythology: The case of the zero-copula. *Nordic Journal of Linguistics* 17: 105–126.
- Underhill, Robert (1976). *Turkish Grammar*. Cambridge, Mass.: MIT Press.
- van Driem, George (1987). *A Grammar of Limbu*. Berlin: Mouton de Gruyter.
- Weber, David J. (1989). *A Grammar of Huallaga (Huánuco) Quechua*. Berkeley: University of California Press.