

Grammaticalization, Typology, and Semantics:

Expanding the Agenda

Zygmunt Frajzyngier

University of Colorado

Zygmunt Frajzyngier

Dept. of Linguistics, Box 295

University of Colorado

Boulder, Colorado 80309

USA

Zygmunt.Frajzyngier@Colorado.Edu

1 Introduction

1.1 The goal and the scope of the paper

The programmatic goal of the present study is to sketch an expanded scope of grammaticalization that covers the emergence of all coding means, including:

- All phonological means
- Lexical categories and subcategories
- Linear order
- Repetition of phrases and lexical items
- Systems of nominal classifications and classifiers
- Inflectional and derivational morphology
- Free grammatical morphemes

The traditional scope of grammaticalization was only the development of grammatical morphemes from free lexical items, hence, covering the last item on the above list, and partially only the next-to-last item, viz. inflectional and derivational morphology. The overlap between the expanded scope and the traditional scope is therefore limited to very narrow areas of the grammar of some languages. Moreover, the expanded scope asks questions that were not asked in the traditional scope, even with respect to free grammatical morphemes and inflectional and derivational morphology. The grammaticalization of the first four means on the list is illustrated in the present study.

Grammaticalization in the expanded sense also includes the emergence of:

- Semantic, pragmatic, and discourse functions-regardless of how these functions are coded within the grammatical system. Traditional studies of grammaticalization deal with the issue of function only as an outcome of the evolution of lexical items into grammatical morphemes.

Grammaticalization in the expanded sense thus covers all processes through which the grammar comes to be. Theoretical questions regarding grammaticalization of means and functions should include:

- Motivations for grammaticalization
- Choice of formal means in grammaticalization
- Consequences of grammaticalization for the grammatical system
- Consequences of grammaticalization for language change

There is no overlap between the theoretical questions of the traditional scope of grammaticalization and the theoretical questions of the proposed expanded scope. The present study attempts to answer these theoretical questions. Grammaticalization as conceived herein provides a powerful tool for explaining language; language change; and for the discovery of functions (meanings) encoded in the lexical categories and in the grammatical system of the language.

1.2 The organization of the present work

The study is organized as follows. In the Introduction I define the terms and comment on the major concerns and criticisms of the traditional approach to grammaticalization. In section 2 the expanded agenda for grammaticalization is fleshed out. Section 3 presents the major theoretical issues and the potential outcomes of the expanded model. Once the questions are formulated, a brief discussion of the possible answers to each question follows. In section 4, I discuss the principle of functional transparency, a major motivating principle for grammaticalization; the principle of indirect means, responsible for grammaticalization in some domains; and the consequences of the initial state for both principles. Section 5 deals with the grammaticalization of selected coding means, viz. phonological means, linear order, repetition of phrases, and the grammaticalization of lexical categories and subcategories. Section 6 is a comparative study of the principle of functional transparency as applied to the coding of grammatical relations in various Chadic languages. Section 7 raises the question of the choice of coding means in

grammaticalization; section 8 deals with the importance of grammaticalization as a motivation for language change; and section 9 points to the use of grammaticalization as a research tool in the study of meaning. Section 10 is a summary of the study.

1.3 The terms

The term ‘**grammar**’ refers to the totality of coding means other than individual lexical items in a given language and to the functions coded by those means. The coding means thus include the phonological means, inflectional and derivational morphology, lexical categories and subcategories, linear order, repetition of lexical categories, reduplication, and all possible combinations thereof. These means, separately or in combination, interact in the coding of various functional domains. One of the characteristics of the interaction of various coding means is their complementarity in the sense that if a given function is coded by one means or a combination of means within an utterance, it is not coded by another means or combination of means within the same utterance. Thus, if the future tense is coded by the auxiliary ‘will’, it is not coded in the same utterance by the structure ‘be going to’. In Polish, the subject is coded through the agreement system on the verb. When a pronominal subject is used in addition to agreement, the pronominal subject codes a different function than agreement, usually the switch-reference function and contrastive focus (Frajzyngier 1997).

The term ‘grammaticalization’ is understood here as the evolutionary process whereby a language develops grammatical means to code various functional domains, whether formal, semantic, or pragmatic. In this usage the term ‘grammaticalization’ is in contrast with ‘lexicalization’, which refers to the development of various lexical means to code the same functions.¹ This extended understanding has its predecessors, in Frajzyngier 1991 and 1996; **Hopper** and **Traugott** 1993; **Giacalone-Ramat** and **Hopper** 1998. Grammaticalization so understood is also akin to the notion of emergent grammar of Hopper 1987. The term ‘grammaticalization’ also includes the development of functional domains and sub-domains as coded by the grammatical means. The notion of functional domains has been used by a number of scholars, e.g. **Givón** 1990. For recent discussions see Frajzyngier and **Mycielski** 1998 and **Croft** 2001.

The prevalent, but by no means unique, understanding of the term ‘grammaticalization’ is confined to the development of grammatical morphemes from lexical items and constructions. This approach has its roots in the works of Wilhelm von Humboldt, later Meillet and Kurylowicz, and more recently Samuels 1972, and it was implemented with respect to a large number of grammatical morphemes across languages by Lehmann 1982 and subsequent editions, Heine and his collaborators in a large number of publications, and a host of other scholars. (For a concise but excellent history of grammaticalization studies in the traditional approach, see Hopper and Traugott 1993.)

The notion of ‘emergent grammar’ refers to Hopper’s conception of the grammatical system as constantly undergoing change. The notion of constant change is not controversial, and the present study is very much in the spirit of Hopper’s ‘emergent grammar’, albeit without some of Hopper’s conclusions. The fact that some grammatical categories are emerging does not imply that there are no categories that have already been grammaticalized.

1.4 Major concerns of the lexical item → grammatical morpheme approach

The purpose of the present section is to consider some theoretical issues that arise with use of the traditional approach. The discussion below also addresses the issues raised by critics of the traditional approach to grammaticalization. The traditional approach is most fully exemplified in the work of Heine and his collaborators, but many other scholars (including this author) work or used to work within this approach.

As recently as 2002, Heine and Kuteva defined grammaticalization as: ‘the development from lexical to grammatical forms and from grammatical to even more grammatical forms’ (p. 2).

They further stated that grammaticalization involves four interrelated mechanisms:

- (a) Desemanticization (or ‘semantic bleaching’)
- (b) Extension (or ‘context generalization’ — use in new concepts)
- (c) Decategorization — loss in morphosyntactic properties of lexical or other less grammaticalized forms

(d) Erosion ('phonetic reduction') —loss in phonetic substance (Heine and Kuteva 2002: 2)

Heine and Kuteva did not define their understanding of the term 'mechanism'. The relevant ordinary English understanding of this term (the term is not a part of standard linguistic terminology as evidenced by it not being included in standard linguistic dictionaries, e.g. Matthews 1997) is 'instrument or process, physical or mental, by which something is done or comes into being (American Heritage Dictionary 1976). The ordinary French understanding of this term is a combination of organs or pieces destined to assure the functioning [of a system]; or 'the ways a system functions' (Larousse de poche 2005: 503). Similar understandings obtain in other languages where this term has been borrowed.

Heine and Kuteva 2002 have explained 'mechanisms of grammaticalization' as the direct or indirect effects of grammaticalization on the lexical sources of grammatical morphemes rather than the means through which the grammaticalization is realized.

A full description of the evolution of a lexical item or a construction into a grammatical morpheme must answer the following questions:

- (a) What is the motivation for the grammaticalization?
- (b) Under what conditions are lexical items rather than other means selected as the sources of grammaticalization?
- (c) What determines the choice of one rather than another lexical item?

Despite the very large number of traditional grammaticalization studies completed within the past 30 years that are devoted to the lexical item → grammatical morpheme change, we have only partial answers to question (c). Heine and Kuteva 2002 is wholly devoted to this issue. The first two questions remain open. Answers to the theoretical questions of the expanded model that will be provided in the present study subsume, however, also answers to questions (a) and (b).

1.5 Criticism of the traditional approach so far

Another issue in the traditional approach to grammaticalization has been the unidirectionality of grammaticalization. Grammaticalization from a lexical item to a grammatical morpheme is certainly more frequent than grammaticalization in the other direction, but there is an explanation for this fact, one based on the lack of motivation for such a process in the latter direction: The shift from grammatical structure to a lexical item would in most cases involve the **lexicalization** of a function that is already coded in the language either by a specific construction or by lexical items that are components of the construction. The creation of another means of coding a function already coded in the language would be an unmotivated language change.² This explanation for the statistical unidirectionality of the lexical item → grammatical morpheme change is preferable to **Haspelmath**'s 2004 defense of unidirectionality in which he pointed out, among other things, that the development of lexical items from grammatical morphemes is non-systematic. But isn't any lexicalization by means other than derivational morphology non-systematic? What is systematic about English 'whey' and **Migaama** (East Chadic) *kàmé* 'hunt mice' (Frajzyngier 2004b)? The relatively rarely attested cases of the change of a grammatical construction into a lexical item bear the same unsystematic characteristics as any other lexicalization, e.g. the jocular but widespread 'druthers', as in 'Will you let me know your druthers' from 'would rather' in English.

Frajzyngier 1996 and 1997 showed that unidirectionality is theoretically unjustified and factually incorrect with respect to one grammatical morpheme changing into another grammatical morpheme. The notions 'more grammatical' and 'less grammatical' and 'more concrete' and 'less concrete' first used by Kurylowicz with respect to lexical items have no known definitions or diagnostic tools with respect to relationships among grammatical morphemes.

Heine, **Claudi**, and **Hünemeyer** 1991 saw communicative needs, problem solving, and creativity as the main motivations for the employment of old means for new functions. Hopper and Traugott 1993 postulated speakers' interactions, communicative processes, and economy as among the motivations for grammaticalization. These motivations are much too vague to provide predictions about grammatical structures and, as Hopper and Traugott noted, may imply the teleological nature of language. None of these motivations can be used as research tools, because every instantiation of language

use is motivated by communicative need, by speaker's interaction, or some communicative problem solving. Cliché and idiomatic expressions apart, every instantiation of language use is also an example of some type of creativity.

Much of the previous work on grammaticalization has been vigorously criticized by Campbell and Janda 2001, Campbell 2001, Joseph 2000 and 2004, and Newmeyer 2001 and his other writings. These writers have pointed out that unidirectionality is not absolute; that there is grammaticalization without semantic bleaching; and that phonological reduction is a result not of grammaticalization but rather of the frequency of use. They also had objections related to the status of grammaticalization among linguistic sub-disciplines. These objections were partially addressed in Haspelmath 2004. All the criticisms of grammaticalization theory and practice pertained, however, only to the lexical item → grammatical morpheme scope, even though for some time studies of grammaticalization with a wider scope have been available.

2 The content of the expanded model: The means and functions

In the present section I flesh out the expanded notion of grammaticalization as a process through which the grammatical means and the grammatical functions come about. The grammar of any language is the product of this process, a product that undergoes continuing change.

The formal means resulting from grammaticalization include:

- **Phonological means**, such as use of tone to code grammatical distinctions; intonation; pauses; vowel reduction and retention; and vowel and consonant harmony
- **Lexical categories** and subcategories (but not individual lexical items), such as but not limited to nouns, verbs, adjectives, auxiliary verbs, prepositions and postpositions, complementizers, determiners, subordinating particles, and intensifiers
- **Linear order** in all of its instantiations, i.e. with different lexical categories in its scope; e.g. use of word order to code questions; grammatical relations; aspects; information structure; and head-modifier relations

- Inflectional means on all lexical categories, e.g. on nouns, verbs, auxiliaries, adjectives, adverbs, adpositions, complementizers, and determiners
- Repetition of lexical items and phrases (as illustrated below)

It is important recognize that none of these means, with the exception of free grammatical morphemes, inflection, and derivation, have inherent or iconic semantic properties that make them particularly appropriate for the grammaticalization of any specific functional domain. Even the repetition of a lexical item or phrase, which is often taken for granted, does not have to have iconic properties. In many Chadic languages, repetition codes, among other things: perfect or perfective aspect; progressive aspect; adverbs; and discourse continuity. Plurality of nouns or plurality of the event is about the only iconic function of repetition. Out of these five coding means, only free grammatical morphemes and some inflectional and derivational morphemes may have lexical items as their source. For many languages, these morphemes represent a relatively small part of the grammatical system. It is this aspect of grammaticalization that has received the most attention, resulting in abundant lists of lexical sources of grammatical morphemes. Given the advances of knowledge in this type of grammaticalization, it will not be elaborated in the present study.

Grammaticalization also includes the emergence of means that do not carry a specific meaning by themselves but that contribute to the creation of other means used to code functional domains. To this group belong:

- Systems of nominal classification (including genders) to enable a system of reference (Frajzyngier and Shay 2003)
- Systems of nominal classifiers to enable a system of individuation, of number coding, and in a few cases also of reference across discourse. (For a different interpretation of both of these systems, see Aikhenvald 2000.)
- Phonological means again, such as pauses, vowel reduction, and retention

The study of the emergence of these coding means requires much work. We know next to nothing about the grammaticalization of phonological means, although we have

very substantial knowledge of the structure of phonological systems. An example of a possible question here is why some languages have grammaticalized vowel retention as a marker of phrasal boundary and other languages have not.

We know from the pioneering work of Kurylowicz, followed by Greenberg and subsequent studies by others, how grammatical gender emerges. But we know very little about the conditions under which a language develops gender systems. The crucial question here is why some languages have developed a gender system and other languages have not. The answer must be that some languages have resolved the issue of the coding of reference by means that do not require gender.

Traditional grammaticalization was often formulated in terms of the source and the target. This conceptualization is applicable only to the evolution from a lexical item to a grammatical morpheme. It has no place when we are talking about the grammaticalization of formal means other inflectional and free morphemes or the grammaticalization of functional domains.

Another facet of grammaticalization is the emergence of functions. The grammaticalization of functions involves the coding of a function that hitherto might not have been available in the language or that might not have been coded by grammatical or lexical means. This latter case occurs when a given function is expressed by periphrastic means potentially different for different speakers.

The types and number of functions constitute an open set and traditionally include various types of modalities; predications; semantic relations between predicates and noun phrases; relationships among noun phrases; pragmatic functions including information structure; discourse functions including backgrounding; and all kinds of relationships between the speaker and the listener. But this is just a selection of a few very general domains. Any non-aprioristic grammar describes a host of functions that have not been noted in other languages or that do not fit within the known functional domains.

The reason I do not identify the grammaticalization of means with the grammaticalization of functions is that once a means is grammaticalized, its function may change. Moreover, once a function has been grammaticalized, it can be coded by a number of means. Studies of language change provide ample evidence for the substitution of coding means and for emergence of new formal means to code the same

function, albeit often with functional extensions or reductions (compare gender coding through nominal endings in Latin and gender coding in French through articles). The grammaticalization of one function may in itself be the motivation for the grammaticalization of another function within the same domain. An example of such a scenario is the grammaticalization of the ‘polite imperative’ after the grammaticalization of the imperative in several Chadic languages. The study of the grammaticalization of functions is in its infancy and awaits both theoretical and methodological developments.

I shall illustrate grammaticalization for the remaining types of coding means and concentrate on two issues: the grammaticalization of lexical categories and linear-order constraints, both as factors motivating further grammaticalizations. A well-argued explanation of the coding means used for one function requires the explanation of other coding means used for other functions within the same domain, a requirement that by far exceeds the limitations of space in the present study. In addition, given the fact that grammaticalizations from sources other than the lexicon have not received much attention, my discussion below contains more questions than answers.

Campbell 2001 challenged the legitimacy of using the term ‘grammaticalization’ in the expanded sense, the sense advocated in the present study. His argument was that the term ‘grammaticalization’, when it applies to the whole grammar, is identical with the terms ‘grammar’ and ‘emergent grammar’. He stated that the term ‘grammar’, ‘emergent grammar’, or ‘typology’ will do (Campbell 2001: 155). Since Campbell did not define his understanding of the term ‘grammar’, his claim that other disciplines cover the same issues as grammaticalization remains quite vague. Behind Campbell’s criticism of the expanded scope of grammaticalization lies a claim that other sub-disciplines within linguistics, presumably descriptive grammar, historical linguistics, and typology, deal with the expanded issues of grammaticalization adequately. Very few descriptive grammars discuss the grammaticalization of either the coding means or the functions in the language under description. And even the grammars that do discuss grammaticalization cannot devote much space to this issue. Traditional historical linguistics could and should, but has not addressed the fundamental questions of grammaticalization as conceived in the present study, viz. the motivation for grammaticalization, the choice of grammatical means, and the consequences of

grammaticalization. One would in vain look for discussion of those issues in Campbell's 2004 *Historical Linguistics*, even though that second edition contains an expanded discussion of grammaticalization in comparison with the first edition. Campbell's objections to the use of the term 'grammaticalization' in the expanded sense can be rejected on two grounds. The first is that no single theoretical approach, methodology, linguist, or a group of linguists has exclusive custody of the term 'grammaticalization'. The second reason for ignoring Campbell's objections is that the expanded scope of the term 'grammaticalization' subsumes the traditional scope of grammaticalization as one of the means by which languages build a grammatical system, rather than **the** means of building a grammatical system.

Typological linguistics, especially typologies of hundreds of languages, moreover, do not address the diachrony of typological phenomena, the motivation for the existence of the coding means, the grammaticalization of meaning, or the change of functions. The problem here is that there are too many potential causes for any single phenomenon. Descriptive grammars, historical linguistics, and typological linguistics do provide, however, necessary data and tools for the study of grammaticalization. Grammaticalization with an expanded scope should become a fundamental part of historical linguistics. Traditional historical linguistics, i.e. the study of language change, typological linguistics, and grammaticalization together, combined with the study of neurological constraints, can explain why languages have the structures they do.

3 Theoretical questions in the expanded model

There is one fundamental question about grammaticalization and two subsidiary questions within the expanded scope of grammaticalization.

The fundamental question in the study of grammaticalization is What are **motivations for grammaticalization**? The traditional assumptions about motivation as listed earlier in this study are much too general to explain differences across languages. Within the expanded model, the question about motivation concerns specific forms and functions rather than the behavior of language in general.

The first subsidiary question relates to the choice of formal means for the grammaticalization of a given **function**. This issue can be divided into several questions: (a) Is there any order in the choice of coding means, i.e., do speakers explore some means first and later on explore other means? (b) Is the choice determined by the functional domain to be coded? and (c) What properties (if any) must a given means have in order to be used for the coding of a given function? For example, why in some languages are polar questions coded by word order and in others by dedicated interrogative particles? We know what the correlates of each means are, but we do not know enough about the choices in particular languages.

The second subsidiary question is What are the consequences of grammaticalization for language use and language change? Once a function has been grammaticalized, it forces different speech behavior on the part of speakers. The speakers must henceforth encode the grammaticalized domain, as per **Jakobson**'s observation that languages differ not in what they may express but in what they must express.

Given the common human anatomic makeup, common cognitive apparatus, common living and cultural concerns, the question why some languages have grammaticalized functions that other languages have not may well be one of the most important and most difficult questions in grammaticalization theory. Many functions were grammaticalized opportunistically, as a by-product of the grammaticalization of other functions. Thus in Gidar (Central Chadic) content questions about human participants in an event, equivalents of 'Who came?' can make a distinction between 'who (feminine)', 'who (masculine)' and 'who (plural)' because question-words code gender and number distinction (Frajzyngier forthcoming). But is this an explanation for the grammaticalization of all functions? Some logicians and even some linguists (e.g., **Wierzbicka**) argue that there exists a set of semantic universals that must be expressed in every language. Even if that were true (and there is no evidence that it is), what about the many functions that are not universal? These theoretical questions are taken up in the ensuing parts of this paper.

Another question in grammaticalization is about the conditions under which a given formal means acquires a new function. One of the conditions under which such a process might occur is that the new function constitutes a metaphorical or metonymic extension

of the function already carried by the formal means, as amply documented by Heine's work with respect to lexical categories. Change in the functions of linear order, from the coding of the pragmatic function 'topic' to the coding of the grammatical function 'subject', would fall into this category, as described by Shibatani 1991. But under what conditions would a formal means that has been grammaticalized for one function carry an unrelated function? Are there any such cases? This is an open question that merits further investigation.

4 The Motivations for grammaticalization

In this study I shall concentrate on three factors that have a role in grammaticalization, but by no means do I want to claim that these are the only motivations involved. One is the principle of functional transparency (Frajzyngier 2004a, Frajzyngier and Shay 2003). Pragmatic motivations, including the principle of indirect means (Frajzyngier and Jirsa 2005), play an important role in the continuing re-grammaticalization of a large number of functions that have already been grammaticalized. An the third one is the initial state of the language at any given time.

4.1 The principle of functional transparency

One of the fundamental principles driving grammaticalization is that of functional transparency, which states that the function of every form in an utterance must be transparent to the hearer. 'Transparency' refers to the functional domains coded in the given language, rather than to the hearers' need to understand the role of an element in reality. The referential meaning of an utterance may be and often is completely obscure to the hearer, but if the role of every element in the utterance is marked through the coding means available in the given language, the utterance satisfies the principle of functional transparency. The principle of functional transparency has in its scope every element in the utterance.

An anonymous reader questions whether the principle of functional transparency is the same as the principle of anti-ambiguity. I am not aware of any 'principle of anti-

ambiguity’, and indeed, ambiguity is widely attested in languages either as a conscious means used in conversations and literature, or as an accidental or necessary outcome of various phonological, morphological, and syntactic constraints. Therefore, the principle of functional transparency is not a ‘principle of anti-ambiguity’. Since other readers may raise similar questions, here is a brief explanation and illustration. Fundamentally, the construction of any utterance is ruled by the principle of functional transparency. The role of every element in the utterance must be transparent to the hearer. The principle of functional transparency can be assured either by the inherent properties of the lexical items used or by external means, which would mark the role of a given item. The principle of functional transparency for English predicts, among other things, the use of prepositions. Every noun phrase other than the subject and object must be marked for its role. One of the means of marking the role of the noun phrase is through the use of the preposition. Consider the following ungrammatical utterance:

*You are not connected Internet.

The utterance is ungrammatical because the English hearer/reader does not know the role of the noun ‘Internet’. The addition of a preposition and a definite article makes the role of the noun ‘Internet’ transparent, even for people who do not know the referential meaning of the noun ‘Internet’.

You are not connected to the Internet.

The addition of a graphic coding means, a period, makes the role of the noun ‘Internet’ transparent (presumably as the institution that signed the message):

‘You are not connected. Internet.’

Compliance with the principle of functional transparency is responsible for the grammaticality of Chomsky’s ‘Colorless green ideas sleep furiously’ because the role of every element in the utterance is transparent through the use of morphological means,

viz. plural coding on the noun, suffix ‘-ly’ coding adverb, suffix ‘-less’ coding the modifier, and through the use of configuration, ‘green’ before ‘ideas’ as a modifier, affix-less form ‘sleep’ after ideas.

For a full discussion of the operation of the system of functional transparency, see Frajzyngier 2004a and Frajzyngier and Shay 2003.

4.2 The principle of indirect means

The principle of indirect means states that whenever the goals of speech involve certain functional domains, indirect rather than direct means of expression are used. Direct means of expression are those that grammatically or lexically encode a given function without dependence on contextual knowledge. Indirect means of expression are those that, although they code function A within the grammatical system of the language, are deployed for the coding of function B, for which the language has already developed other coding means. The functional domains involved are speaker-hearer relations: At the level of grammar, the principle of indirect means operates in the domain of interpersonal relations, such as: deontic and epistemic modality; negation; content and polar interrogatives; reference systems; and forms of address.

The principle of indirect means starts operating after some means have been grammaticalized to code a given function and have thus become the direct means of coding. Languages vary with respect to the scope of the principle of indirect means, and it is theoretically possible that there are languages in which this principle does not operate. Moreover, within the same language, various social groups may have different norms with respect to the scope of the principle. In many languages, interrogative constructions are used to code deontic modality, e.g. ‘Can you pass the salt, please?’. The principle of indirect means is responsible for the abundance of formal coding means in the domains that are in its scope. Thus in many languages there are a large number of means to code deontic modality, e.g. modal auxiliaries, inflectional coding on the verb, modal adverbs, in comparison to the means to code tense, a functional domain that is usually not in the scope of the principle of indirect means. In English the means of coding

deontic modality include the auxiliaries ‘must’, ‘need to’, ‘have to’, ‘ought to’, ‘should’, ‘shall’, ‘will’, etc., whereas the past tense is coded by inflectional coding on the verb. For a full discussion see Frajzyngier and Jirsa 2005.

4.3 The initial state

The principle of functional transparency and the principle of indirect means are intimately connected with the initial state of the language. The initial state does not refer to a state in a protolanguage. The term ‘initial state’ refers to the state of the language at any given time with respect to two areas: (a) the coding of some functional domain and (b) the availability of coding means. A coding means can be conceived of as a formal niche. If a given formal niche is not used to code a function, it is available as a coding means. If such a formal niche is already occupied by a function, it is not available as a coding means for another function within the same formal or functional domain (Frajzyngier, Krech, Mirzayan 2002).

The functional properties of the initial state pertain to the functional domains and functional sub-domains that are coded at any given time by all the coding means available, i.e. the lexicon, phonological means, linear order, inflectional and derivational means, and lexical repetition. One of the important factors of formal means, neglected in the study of grammaticalization so far, is that the functional properties of lexical items having the same reference may differ substantially across languages. Thus the fact that in English most singular nouns must be preceded by an article, but not in Polish, may be due to the inherent properties of nouns in the two languages (on properties of nouns cf. also Rijkhof 2002: 50-59).

The motivations for grammaticalization, viz. the principle of functional transparency, principle of indirect means, economy in the sense of coding the same function over many instantiations in discourse, all take place within the following conditions:

- (a) The initial state with respect to the functional domains coded
- (b) The initial state with respect to the formal niches already occupied
- (c) The degree of stability with respect to the functions coded

The information about the initial state must also include information as to whether that state is stable or unstable (for one understanding of the notion ‘stable’, cf. Samuels 1972). An unstable system is one in which given functional distinctions are neutralized in significantly extensive environments. An example of such neutralization would be an increased lack of distinction between subject and object and the neutralization of person and number distinctions resulting from the collapse of the inflectional coding on nouns, articles, adjectives, and verbs in English, which in turn was provoked by rather simple phonological changes.

Once a form has been grammaticalized, it may create an unstable system in some parts of the grammar. Thus, the grammaticalization of linear order as a coding means for grammatical relations largely constrains the use of linear order for the coding of information structure and other pragmatic functions.

5 The grammaticalization of means

5.1 The grammaticalization of phonological means

We know how the systems of underlying segments change: by either the addition or the reduction of segments. We also know how their realizations change. We often know what the causes of these changes are, e.g. changes in stress, tone, intonation patterns, and borrowings. We have some very precise hypotheses concerning the mechanisms of those changes (Blevins 2005). The fundamental questions as to why languages differ with respect to syllable structures, phonotactics rules, vowel and consonant harmony, and rhythmic structure of the utterance have yet to be explained. These constraints and rules are relevant for the understanding of larger structures, hence, for grammaticalization theory. There exist phonological rules whose role is to indicate whether a given constituent is part of a larger construction. An illustration of this type of rule is the phonological reduction responsible for the formation of the construct state in Semitic and other Afroasiatic nouns:

- (1) a. *anfey* *ha-ec*
 branches the-tree
- b. *ha-anafim* *shel* *ha-ec*
 the-branches of the-tree
 both: ‘the tree’s branches (roughly)’ (Hebrew, as cited in Heller, n.d.)

There exist rules marking phrasal boundaries. An illustration of such rules is the phrase-internal vowel reduction and phrase-final vowel retention in Chadic. In some Chadic languages these rules have only morpheme-final vowels in their scope; in other languages they also have morpheme-internal vowels in their scope. In the following example from **Mina** (Central Chadic), the noun *trá* ‘month’ is reduced to *tr* (a high central vowel is subsequently inserted for syllabification) to indicate that it constitutes part of the construction with the next item:

- (2) *tór* *láy* *tó* *mítǎš*
 month time GEN hunger
 ‘the year of the hunger’ (Frajzyngier, **Johnston** with **Edwards**
 2005)

In **Giziga** (Central Chadic), internal vowels are reduced in the phrase-internal position:

- (3) *nány* *à* *ɓ* *à* *ʔkàdʔ*, *à* *ʔkàdʔ* *tàtìi* *tá*
 3.IP 3 IRR 3 **get up** 3 **get up** INTENS NEG
 ‘She tried to get up, but she couldn’t get up at all.’
- (4) *à* *rô* *á* *lm-á* *mándâf*
 3 go 3 find-3 **hare**
 ‘She went and found the hare.’

mándáf à *ɓ-áa* *màdàngòs*: *'kál kál à má kâ?*

hare 3SG say old woman 2 run PREP what Q

‘The hare said to the old woman, “Where are you running?”’ (Shay, Grammar of Giziga, in progress, Frajzyngier field notes)

The grammaticalization of word-final and word-internal vowel reduction probably follows a path similar to the grammaticalization of apophony. First the word-final vowels are reduced when the word is followed by another word within the same phrase. Words that do not have a final vowel undergo internal vowel reduction. Then the phrase-final vowel retention acquires a function of its own, that of indicating phrasal boundary. The grammaticalization of the vowel retention as a grammatical marker at this stage is complete.

The grammaticalization of tone as a coding means is another instance not motivated by any inherent semantic properties of the coding means. In Hdi (Central Chadic) the functions of prepositions *da* and *ta* depend on the tone these markers have:

(5) *tà* ‘stative locative preposition’

tá ‘second argument marker’

dà ‘directional locative proposition with non-toponyms’

dá ‘directional locative proposition with toponyms and purpose subordinator’

(6) *lá-b-l-ì* *dá* *xdi*

go-OUT-go-1SG PREP Hdi

‘I went to Hdi.’ (place name)

lá-b-là *dá* *bàtà* *mákwà*

go-OUT-go PREP woo girl

‘He went to woo a girl.’

Compare complements that are not toponyms, and with which the preposition has low tone:

(7) *láb-li dà màkàràntà*
 go-OUT-go PREP school
 ‘I went to school.’

láb-li dà Mbitsa
 go-OUT-go PREP Mbitsa
 ‘I went to Mbitsa (proper name).’

Tone distinctions are exploited to code semantic functions, and their grammaticalization is not linked to any specific semantic function to be grammaticalized or to any inherent semantic properties of high or low tone.

5.2 The grammaticalization of linear order

Linear order is a readily available means to code a relationship between any two entities, whether they belong to the same or to different lexical categories. When linear order involves lexical items belonging to the same category, one function is assigned to the first element and another function to the second element. This is the situation when one noun modifies another in English. One noun is the head and the other is the modifier. When linear order is applied to different lexical categories, the function of only one element is coded by the order, the function of the other being that of the reference point, determined probably by its categoriality. Thus in a VS language, the function of the noun as the subject is coded by the position after the verb. The function of the verb as the predicate is not coded by the position before the subject but by some other means, presumably morphological.

Linear order as a sole coding means inherently has very limited possibilities. First, it requires a point of reference (for a discussion see Frajzyngier and Shay 2003). With lexical items belonging to the same categories, linear order can code one functional sub-domain, e.g. modification, when the modification relates to two nouns. When it concerns the relationship between the verb and its arguments, linear order can code at most two

functions, one by the position preceding the verb and another by the position following the verb. One of the open questions with respect to grammaticalization is why some languages are verb initial, some are verb final, and some are verb medial. Given the implications of the position of the verb for the other elements of the grammar, this remains one of the key questions for the understanding of grammatical systems.

5.3 Repetition

Repetition as a coding means is different from reduplication, which involves contiguous repetition of a lexical item or part of a lexical item. Repetition does not have to be contiguous. It may have various lexical categories in its scope, i.e., it may involve the repetition of nouns, pronouns, verbs, and adjectives, and it may involve the repetition of whole phrases. Across languages, the repetition of a lexical item may be used to code different functions. Contrary to frequent tacit assumptions, the repetition of lexical items does not have to be an iconic means. Since this category has not been isolated as an independent coding means so far, the present study illustrates the mechanisms and functions of repetition, and in one case, a motivation for its existence.

In **Wandala** (Central Chadic, spoken in Northern Cameroon and North Eastern Nigeria), a pronoun in the subject function is deployed twice in the sentence, once before the verb and once after the verb. The function of this repetition is not yet clear, but the same means is used to code focus on a pronominal subject:

- (8) *šóì ɣánnà kà cìn-ná-kà*
 story DEM **2SG** listen-3SG-2SG
 ‘The story that you heard’ (Wandala, Frajzyngier field notes)

In **Gidar** (Central Chadic, Northern Cameroon), the verb is repeated at the beginning of a clause, sometimes with its object, sometimes without it, to code focus on the predicate:

- (9) *ilèṅ à-dàw lèṅé-ṅ káká á*
 sleep 3M-G.PROG sleep-PL only
 ‘They are asleep.’

The answer to the question ‘what did he do with this thing [visible]’ is:

- (10) *áddò à-dó-nì*
 reheat 3M-reheat-3M
 ‘He reheated it.’

Non-focused expressions have no repetition of the verb:

- (11) *à-dóhò-ná-n-kò*
 3M-reheat-3M-PL-PRF
 ‘They reheated it.’ (Frajzyngier forthcoming)

Cognate objects, which require a formation of a verbal noun in a number of languages (Goldenberg 1971), most likely represent the same type of grammaticalization, albeit used for different purposes.

A yet-different function of the repetition of a lexical item occurs in **Kafa** (Omotic). Cognate objects are derived from verbs, and thus they represent an instantiation of a repetition. Transitive verbs in this language appear to require an object. If no object occurs, a cognate object is used instead:

- (12) *bóonò dóocò ùcyétè bóonò búunò ùcyétè*
 they beer drink:3PL they coffee drink:3PL
bóonò máyò mà-hé-tè
 they food eat-PRES-3PL
 ‘They drink beer, drink coffee, eat the food.’

‘ínòcé šùnòc’í mmó nà
 when work finish-give ASSC
yùbbò yùbbè ‘í dùubó dùbbè ‘í šòddò šòddè ‘ì nòcé
 choir singing sing sing responsive singing while
 ‘When the work is finished, they sing (choir-like style), sing, sing in a
 responsive style . . .’ (Z.F. field notes)

5.4 The grammaticalization of lexical categories

The grammaticalization of the major lexical categories, nouns, verbs, adverbs, and adjectives, has been largely neglected in grammaticalization studies, and yet whether a language has or does not have a given lexical category may determine the formal characteristics of clauses and sentences in the language. In contrast to the so-called major lexical categories, the grammaticalization of adpositions, complementizers, and subordinating particles has received ample attention and is relatively well understood as far as the sources of these morphemes are concerned. The present study sketches only how one might study the grammaticalization of the major lexical categories, and it provides a principled answer as to why certain languages have some categories rather than others. Once the grammaticalization of lexical categories is understood, one can place in proper perspective the grammaticalization of derivational morphemes, another area that has been relatively well developed in grammaticalization studies so far.

5.4.1 The grammaticalization of nouns, verbs, and adjectives

As demonstrated in Frajzyngier and Shay 2003, the existence of the lexical categories ‘verb’ and ‘noun’ is motivated by propositional predication, which requires the identification of a predicate and of arguments. But this general requirement can be satisfied in a number of ways, not necessarily through the creation of lexical categories. One of these ways is the use of free grammatical morphemes coding the categories

‘tense’ and ‘aspect’, whose presence in the utterance indicates that an accompanying lexical item is a predicate, and in South East Asian languages, the existence of classifiers, whose presence in the utterance indicates that the accompanying lexical item is an argument.

A modifying function can be coded by the lexical category ‘adjective’, but it also can be coded by constructions akin to relative clauses and by a number of other means (Frajzyngier and Shay 2003). We can only speculate at this stage about the mechanisms for the grammaticalization of lexical categories. One such process would be as follows: A lexical item undifferentiated for a lexical category would be accompanied by other items, such as determiners, pronouns, time adverbials, and quantifiers, to code those functions, thereby identifying the lexical item as a predicate or an argument. These lexical items eventually fuse with the items they were modifying, creating the categories of gender-marked nouns as per the scenario conceptualized by Kurylowicz and elaborated by (Greenberg 1978) and tense or aspect coding on verbs. Mandarin would represent a language where individual lexical items do not have an inherent categoriality. Predicate and argument functions are coded by isolated grammatical morphemes. I shall not pursue this issue in depth in the present study. Instead I would like to concentrate on the grammaticalization of a lexical category that has not been discussed in the literature other than in my own work on Mina.

5.4.2 *The grammaticalization of the locative predicator*

The grammaticalization of the lexical category ‘locative predicator’ is interesting for a number of reasons. The first is that we can describe, document, and provide the proof of how the lexical category is grammaticalized. We can also provide a motivation for this grammaticalization. Based on this grammaticalization, we can also provide a fine-grained semantic analysis of nouns and verbs and provide the evidence for the proposed subcategories of nouns and verbs.

Mina, a Central Chadic language, has the category ‘locative predication’, which could be either stative (the presence of an entity or an activity at some place) or dynamic

(movement toward or from a place). The form of locative predication in Mina depends on the semantic feature LOCATIVE in the predicate and in the intended locative complement. The semantic feature LOCATIVE is present in directional verbs of movement and in the verb corresponding to ‘be in place’. The semantic feature LOCATIVE in nouns is present in toponyms and in a few other nouns, such as ‘room’, ‘hut’, ‘home’, and ‘compound’. When both the predicate and the locative complement have the semantic feature LOCATIVE, locative predication consists of the simple juxtaposition of the predicate and the locative noun (all data are from natural discourse, as presented in Frajzyngier, Johnston with Edwards [2005], and Frajzyngier and Shay 2003).

The verb *tsú* ‘go to’ is inherently locative, as is the noun *dámù* ‘bush’. Consequently, no additional predicators or prepositions are used in locative predication:

- (13) *séy m̀̀ ng̀̀l ng̀̀l t̀̀y*
 so REL husband husband see
á t̀̀y-ú ẁ̀l tsú z̀̀ d̀̀m̀̀
 3SG see-3SG wife went EE bush
 ‘So the husband saw that the wife went to the bush.’ (EE end of episode)

If the complement is not inherently locative, locative predication requires the use of the preposition *n*, as in the following example where the preposition precedes the noun *hìd̀̀* ‘man’. (Note that in the first line the verb *táy* ‘return’ is inherently locative; hence, no additional predicator is needed, and the noun *wùtá* ‘compound’ is also inherently locative, hence, no preposition):

- (14) *tséy hìdì wàcíŋ táŋ z wùtá à ǰá*
 so man DEM return EE house 3SG say
á n médiŋ ngàŋ wàcíŋ ngámbù
 PRED PREP neighbor 3SG DEM friend
há kà déb-é-ŋ dál nà
 2SG INF bring-GO-3SG money PREP
hìdà wà dál vání
 man DEM money how much

‘When the man came back to the house, he said to his neighbor, “Friend, you brought money to this man. How much money?”’

The absence of the feature LOCATIVE in nouns constitutes the motivation for the emergence of the locative preposition, whose function is to code the grammatical role of the noun phrase rather than a spatial configuration, a function that is coded by other markers.

When the predicate is not inherently locative, locative predication requires the use of the locative predicator *a*. The verb *dáhà* ‘exist’ is not inherently locative; hence, the locative predicator is required in locative predication. The complement noun *bíŋ* ‘house’ is inherently locative; hence, there is no locative preposition:

- (15) *háǰàm dáhà á bíŋ ngàŋ*
 girl exist PRED house 3SG
 ‘There is a girl at her house.’ (*háǰàmà* ‘girl’)

The verb *déb* ‘take’ is not inherently locative, but the noun *ìdá* ‘home’ is. Hence, locative predication requires only the predicator *á*:

- (16) *séy báy déb zà ngàn ká á idá*
 then chief take EE 3SG POS PRED home
 ‘Then the chief took the calabash home.’

When neither the predicate nor the complement is inherently locative, locative predication requires both the locative predicator *a* and the locative preposition *n*:

- (17) *séy hidì mà sà tápá dáhà á*
 so man REL drink tobacco exist PRED
nà mǎŋ màcǐŋ
 PREP ANAPH DEM
 ‘So there is a smoker among them.’

The analysis of locative predication in Mina has the following implication for the general theory of language and for grammaticalization theory. The grammaticalization of lexical categories and grammatical markers, in this case the locative predicator *a* and the locative preposition *n*, is motivated by the principle of functional transparency, as discussed in the next section in the present study.

6 A case study: The principle of functional transparency and **grammatical relations**

In the present study, the principle of functional transparency is illustrated on the relative clause construction in **Wandala**, a Central Chadic language. This discussion also serves to illustrate the grammaticalization of the repetition of a lexical item as described above.

6.1 Relative clauses in Wandala

Wandala clauses have subject pronouns obligatorily preceding the verb, and nominal subjects or nominal objects following the verb. The head of the relative clause precedes the relative clause. The relativization of the subject has the form: NP (DEM) *a* (subject pronoun) verb (NP[object]). (The first line in Wandala provides a broad phonetic transcription with spaces indicating short pauses. The second line provides the underlying morphemes):

- (18) *á m̀ gdz̀ yánnà à blà lác nà*
 PRED PLACE child DEM 3SG sing song DEF
sà-wà
 arrive-VENT
 ‘Where is this child who sings the song, let him come!’

The relativization of the object has the form: NP (DEM) --subject pronoun--verb--object pronoun--subject pronoun, i.e., the subject pronouns both precede and follow the verb:

- (19) *názù yè kàtà ndà-ká-yà vá-t̀-ñ wá*
 thing 1SG want say-2SG -1SG time-PREP-DEF COMP
 ‘What I would like to tell you today is . . .’

If the subject of a relative clause with a relativized object is nominal rather than pronominal, the third-person subject pronoun precedes the verb and the nominal subject follows the verb:

- (20) *názù àfámyàrmàlámì nákátáfká mínà*
názù à fá-myà-r màl-á-mínà kà-t
 thing 3SG put-1PL.INCL-ON boss-GEN-1PL.INCL PREP-PREP
áfík-á mínà
 front-GEN 1PL.INCL
 ‘that what has been asked by our boss who is in front of us . . .’

Although the position of the nominal subject after the verb is the default position for this category, the position of the pronominal subject after the verb is not, at least not for the great majority of verbs. Therefore, the position of subject pronouns after the verb in relative clauses needs to be explained. I propose that the function of the subject pronouns occurring after the verb is to indicate that the noun that is the head of the relative clause is not interpreted as the subject of the clause. Such an explanation follows from the principle of functional transparency. If the head of the relative clause is the subject, the subject pronoun is not repeated (example (21) is the same as ex. (18)):

- (21) *á m̀ gdz̀ òyánnà à ̀l̀ l̀x̀ ǹ*
 PRED PLACE child DEM 3SG sing song DEF
s̀-ẁ
 arrive-VENT
 ‘Where is this child who sings the song, let him come!’

Interestingly, the principle of functional transparency is not universally required with respect to relative clauses, as there are languages, e.g. some Semitic languages, where relative clauses are ambiguous with respect to the grammatical role of their heads (cf. Frajzyngier and Shay 2003, with reference to Goldenberg, p.c.). Why it is that in relative clauses across languages the distinction between the subject and the object may be neutralized merits further exploration. Tentatively, I propose that the relative clause is merely a clausal comment on the head. In many languages the head does not have to be an argument of the relative clause, and therefore the question about the role of the head of the relative clause is not within the scope of the principle of functional transparency. So

even if the head happens to be the argument of the relative clause, the role of such argument is not overtly coded.

6.2 Coding grammatical relations in a simple clause

6.2.1 *Word order and the coding of grammatical relations*

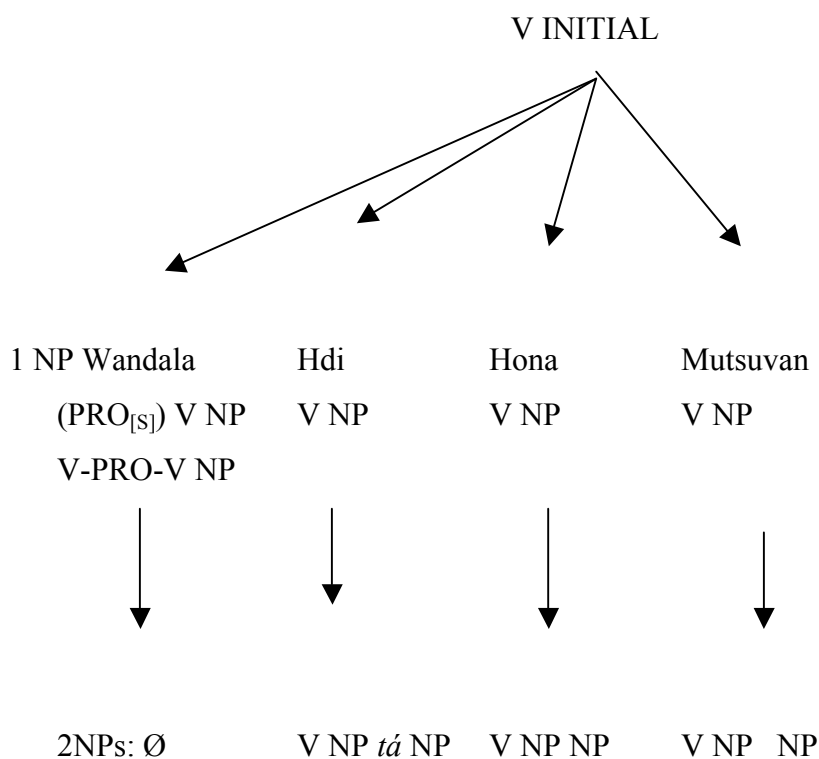
The discussion of coding grammatical relations in a simple clause aims to examine the interaction of the characteristics of the initial state, the principle of functional transparency, and the choice of grammatical means. In several Central Chadic languages the verb is in the clause-initial position. This position of the verb constrains the coding of the role of the argument by position to one argument only: V NP. The argument occupying such a position can be associated with one grammatical relation, e.g. subject or object, categories coded in many Chadic languages, or with one semantic role, categories also coded in some languages with respect to the second argument. The presence of the verb in the clause-initial position posits the following issues:

- (a) How is a second argument added to a proposition?
- (b) How is the grammatical role of the single argument of the transitive verb coded? This is not a trivial question, because the single argument may be the outcome of the omission of another argument, either the subject or the object.
- (c) If a verb can occur with two arguments, what happens if one of the arguments is fronted? How is its role coded, and how is the role of the remaining argument coded?

Questions (b) and (c) are likely to have the same answers in each individual language. All of these questions are issues for grammaticalization, i.e. for the development of coding means for various functions.

Chadic languages with the verb in the initial position display a variety of strategies for the coding of an additional argument. Each strategy is an instantiation of a different grammaticalization, and that is why they deserve to be discussed in this study. Chadic languages have also developed different strategies for the coding of the roles of the single argument in cases of omission and fronting. None of these strategies involves a lexical item → grammatical morpheme process. The initial state with verb-initial position is discussed for four languages: **Hona** (spoken in Nigeria, data Frajzyngier field notes, and

Jordan field notes), **Hdi** (Northern Cameroon, Frajzyngier with Shay 2002), Wandala (North Eastern Nigeria and Northern Cameroon, Frajzyngier field notes), and **Mutsuvan** (Northern Cameroon, **Johnston** p.c., from his field notes). Each of these languages has solved the issue of adding one more argument to the clause in a different way, despite the fact that the languages are genetically related, belong to the same branch within the family, and are spoken within the same geographical area. Crucially, however, none of these languages is an immediate neighbor of any of the other three languages. Hdi is verb-initial in all aspects in pragmatically neutral clauses. Wandala has pronominal subjects preceding the verb, but all nominal arguments follow the verb. Hona is verb-initial in the perfective aspect and verb-medial in the imperfective aspect. Mutsuvan appears to be verb-initial in all pragmatically neutral clauses. Assuming verb-initial to be the initial state, the four languages have grammaticalized the following strategies for the coding of one argument and distinguishing between the subject and the object functions of two noun phrases:



Hdi and Wandala have reduplicated forms of the verb in pragmatically independent clauses. In Wandala, the subject pronoun is inserted between the reduplicated forms, and in Hdi subject pronouns are suffixed to the reduplicated form of the verb.

With intransitive verbs, Hdi, Hona, and Mutsuvan have the structure V NP. There are no structural distinctions between Hona and Mutsuvan at the two-nominal-argument configurations. The difference between the two languages emerges in other structures, viz. when one of the arguments is omitted or when one of the arguments is fronted. What one would like to know is whether there is a principled reason for why the four genetically related languages, starting from what appears to be the same initial state, have grammaticalized different strategies for the coding of subject and object. In what follows I describe the inclusion of two arguments, the grammatical role of the single argument, and the coding of the roles when one of the arguments is fronted (given the fact that the languages are verb-initial, placement of an argument in the clause-final position does not exist as a coding means).

6.2.2 *Two arguments in a pragmatically neutral clause*

The simplest means to code a relationship between two elements is by juxtaposition. This means is deployed in Hona and in Mutsuvan. In Hona, the nominal subject follows the verb, and the nominal object follows the subject. The verb does not have any pronominal arguments. Hona does not use prepositions to code grammatical relations (all examples in Hona are elicited):

- (21) *fí* *nú-nà* *ʔú-wà*
 bake wife-1SG meat-INDEF
 ‘My wife baked meat.’

Pronominal subjects are suffixed to the verb. The nominal object follows the verbal complex. The noun phrase that follows the verb is interpreted as the object because the subject is coded by the suffix on the verb:

- (22) *kyàtlá-ndà* *dèfí-nà-d* *pat*
 cut-3PL tree-DEF-ALL all
 'They cut all the trees.'

Pronominal objects are also suffixed to the verb. The third-person singular subject pronoun is formally different from the third-person singular object pronoun. If the clause has a nominal object, the pronominal object suffix codes the definiteness of the object:

- (23) *ʔe-n* *Ali* *hur-ghen*
 slaughter-OBJ Ali neck-3SG
 'Ali slaughtered it.' (lit. 'Ali cut its neck')

The question remains why the simple juxtaposition we have seen in Hona and Mutsuvan is not deployed in Wandala and Hdi. Each of these languages presents a different case.

Hdi

The means used for coding grammatical relations in Hdi depend on the aspect and the information structure and other pragmatic functions coded in the clause. There are two aspectual systems, one coding pragmatically dependent and the other, pragmatically independent clauses. Pronominal subjects are suffixed to the simple or the reduplicated form of the verb. The third-person singular subject is unmarked in the perfective, and in some types of clauses it is marked in the imperfective. If there is a nominal subject, no pronominal subject is added to the verb. Pronominal objects are infixes between the reduplicated form of the verb. The third-person pronominal direct object for inanimate nouns is unmarked. The nominal object can be marked by the position following the verb if the subject precedes the verb and is marked for its role. This is the case in sequential clauses:

(24) *dg-áy-tán tá dgú yá*
 thresh-PO-3PL OBJ threshing DEM
 ‘While they were threshing’

mbàd' ká pákává ghívì kà klà-á-tá vərə
 then COMP hyena SEQ take-PART-REF beans
 ‘Hyena took some beans.’

If the nominal subject follows the verb, the nominal object, whether direct or indirect, is marked by the preposition *tá*, derived most likely from the locative stative preposition *tà* through tone-raising:

(25a) *ngá dà-gá-ghà-tà índà grá-xà-ní*
 NORM cook-INN-D:PVG-REF:SUBJ all friend-PL-3SG
tá ghzú
 OBJ beer
 ‘All of his friends should cook beer and bring it there.’

(25b) *kà zì-í-n-tá tsá mighám yá [tá*
 SEQ chase-AWAY-3-REF DEF chief DEM OBJ
zón-á-ní] tá tsá màrkw-á-tàn yá
 child-GEN-3SG [error] OBJ DEF wife-GEN-3PL DEM
 ‘The chief chased away his wife.’

Pronominal subjects are suffixed to the verb, and the second argument that follows the verbal complex is marked by the preposition *tá*, just as in the case when the nominal object follows the nominal subject. The reason that Hdi codes the second argument by the preposition is that otherwise the role of the second noun phrase would not be interpretable as an argument of the verb.

Thus, if the nominal argument after the verbal complex were to be unmarked for its role, it could be interpreted as the subject. The grammaticalization of the preposition to code the object is thus motivated by the principle of functional transparency.

Wandala

The reason that two arguments cannot occur after the verb in Wandala has to do with the fact that the role of the single argument following the verb is unmarked and must be calculated from the coding of other elements. The verb *s* ‘arrive, come’ can be intransitive or transitive, and the nominal argument following it can be the intransitive subject, transitive subject, or object, depending on the inflectional coding on the verb:

- (26) éməlmàʒárà sàrəmsámdə̀nà
 é màʒárà sá-rə-m-s-à mdə̀ ná
 well now arrive-3PL-IN-arrive-GO people DEF
 ‘And now, those people came.’

The nominal argument following the verb is interpreted as the object if the pronominal subject has a different number than the nominal argument:

- (27) tàdə̀sə̀sá dàdámúvgè
 tə̀ də̀ sə̀-s-á dàdà á-m úvgè
 3PL SEQ arrive-arrive-GO father PRED-IN grave
 ‘And they raised the father from the grave.’

Putting two nominal arguments after the verb would leave the clause uninterpretable with respect to the distinction between the two arguments. One means to include two arguments in the proposition is to deploy one clause with an intransitive verb and a nominal subject and then reference that subject by a pronoun in the subsequent clause with another nominal argument. The verb *tsá* ‘get up’ and equivalents of the verbs ‘go’ and ‘come’ are used in the introductory clauses (all data from natural discourse):

- (28) tátsámdá žílé
tá tsá md-á žílé
 3PL get up people-GEN man
 ‘People of the groom get up.’

tásə̀bə̀dʒá mdámùksé
tá sə̀ bə̀dʒá md-á mùksé
 3PLarrive meet people-GEN woman
 ‘They come to find the people of the bride.’

Another means to include two arguments is the fronting of one of the arguments, mainly for topicalization. The grammatical role of such an argument is computed from the inflectional coding on the verb and from the presence of other arguments. In the following example, the evidence that the noun *žílé* ‘man’ is the topic is provided by the fact that it retains the word-final vowel. Its interpretation as the subject is assured by the presence of the goal marker on the verb *d-á* ‘do-GO’ = ‘bring’:

- (29) žílé ádá názù ádə̀ndə̀və̀nánə̀nə̀né
žílé á d-á názù à də̀ ndə̀-v-ə̀n-án ɲáne
 man 3SG go-GO that 3SG go ask-AFF-3SG-ASSC 3SG
 ‘The man should bring that with which he is going to ask for [the hand of the girl].’

It appears that the only cases in Wandala where a clause has two nominal arguments in natural discourse occur when one of the arguments is topicalized or focused.

Summary of the coding of clauses with two arguments

	Wandala	Hona	Hdi	Mutsuvan
Two arguments	∅ (or S[<small>TOP</small>]VO), O[<small>TOP</small>] VS	VSO	VS <i>tá</i> O	VSO

6.2.3 *The role of the single argument*

The determination of the role of the single argument following the verb is interesting for verbs that can occur with two arguments. Most grammars of Chadic languages do not recognize a passive form of the verb or a passive construction. The default role of the single argument of potentially transitive verbs depends on the category of the verb.

Wandala

Wandala has grammaticalized several means that interact with the inherent properties of the verb to code the distinction between the subject and the object.

In the perfective aspect the noun phrase that follows a transitive verb without any markers is the object:

- (30) à *jjə* *zàhè*
 3SG kill snake
 ‘He killed a snake.’

The addition of the goal marker or object marker to such verbs indicates that the following noun phrase is the controlling participant:

- (31a) à *jj-á* *zàhè*
 3SG kill-GO snake
 ‘A snake killed him/it.’

- (31b) *à* *jj-á-trá* *zàhè*
 3SG kill-GO-3PL snake
 ‘A snake killed them.’

For verbs whose inherent meaning does not imply either the controlling or the affected role of the single participant, the interpretation of the single noun phrase as subject or object is computed from the properties of the noun phrase, such as [+human] vs. [-human]:

- (32a) *à* *hàrd-á* *zàrvà*
 3SG farm-IMPF sesame
 ‘He planted sesame.’ (and not ‘The sesame farmed.’)

- (32b) *à* *hàrd-á* *kèllú*
 3SG farm-IMPF Kellu
 ‘Kellu farmed.’ (and not the nonsensical ‘He farmed Kellu.’)

With one class of verbs, the noun phrase that follows the verb is the object:

- (33) *à* *šib-à* *ógdzrè*
 3 SG hide-IMPF child
 ‘He hid a child.’

In order to code the role of the noun phrase as the controlling participant, the verb, whether in the simple or the reduplicated form, must have an object pronoun:

- (34a) *à* *šibà-ná* *ógdzrè*
 3SG hide-3SG child
 ‘The child hid it.’

- (34b) *šìb-á-n-vá-šìbà* *ágdzrè*
 hide-GO-3SG-AFF-hide child
 ‘The child hid it.’

The absence of the goal or the object marker on the verb indicates that the following noun phrase is affected, without implying that there was another agent that performed the action:

- (35) *šìbé-v-šìbè* *ágdzrè*
 hide-AFF-hide child
 ‘The child hid.’

To summarize: The facts of Wandala demonstrate that grammaticalization is motivated by the principle of functional transparency, whereby the role of every element in the utterance must be transparent to the listener. If the role is transparent because of the nature of the lexical items deployed, no other coding means are necessary.

Hdi

The single noun phrase following the verb is marked as the object by the preposition *tá*, the same way as the second nominal argument in a clause:

- (36) *ngázá tá tsá vákú yá*
 NORM eat OBJDEM year DEM
 ‘He should wait out this year.’

The reason Hdi codes the object role of the single argument by a preposition is that the verb does not code the role of the arguments.

Hona

In Hona, the semantic role of the single argument is determined by the inherent properties of the verb. The single argument of the verb *fí* ‘bake’ is the affected object:

- (37) *fì* *lú-dì*
 bake meat-DET
 'The meat baked.'

A noun that cannot plausibly be the affected argument may not be used as the single argument of such a verb:

- (38) **fì* *núnà*
 bake wife
 for 'my wife baked'

Compare the subject coding with intransitive verbs:

- (39) *wúd* *íyá* *bà* *à* *húkàsù*
 return mother VENT PREP market
 'My mother returned from the market.'

A single nominal argument is interpreted as the object if the verb has a pronominal subject suffix:

- (40) *sán-d'* *ám-dì* *pát*
 drink-3SG water-DET all
 'He drank all the water.'

If the verb has an object suffix, the noun phrase following the verb is interpreted as the subject:

- (41) *ʔe-n* *Ali*
 slaughter-3SG Ali
 ‘Ali slaughtered’

6.2.4 Coding the role of arguments in fronting operations

What means, if any, did the four languages grammaticalize to code the grammatical relations when one of the arguments is fronted for whatever pragmatic function? Despite some similarities, there are interesting differences among the four languages.

The fronted noun phrase in Wandala is interpreted as the subject if the noun that follows the verb is coded as the object through the absence of a goal marker and of an object pronoun on the verb:

- (42) *dàdá* *átàdzgwánkà* *kwàrhàrkà*
dàdá *à* *tà* *dzgwá-n* *kà* ***kwàr*** *hár* *kà*
 father 3SG NEG be able-3SG NEG order home NEG
 ‘The father cannot order his family, can he?’

The fronted argument is interpreted as the object if the noun following the verb is coded as the subject through the presence of an object pronoun on the verb:

- (43) *éemákínnì* *dàdà* *ákwàrná* *màmà*
ée *má* *kínnì* *dàdà* *á* ***kwàr-ná*** *màmà*
 well, HYP BCKG father 3SG order-3SG wife
 ‘If it is like that, the father is ordered by the mother.’

In Hdi, the subject function of the fronted noun phrase is computed from the object coding of the noun phrase that follows the verb. The object is coded by the marker *tá*:

- (44) *tsá Gulu yá ná yà-yà*
 DEF Gulu DEM DEM beget-beget
tá zwàni ndá ndghà
 OBJ child:PL ASSC many
 ‘That Gulu begot many children.’

The fronted noun phrase is interpreted as the object if the noun phrase that follows the verb is not marked by the object marker *tá*, i.e., if it is unmarked and hence must be interpreted as the subject:

- (45a) *ndá ndghà zwàni yà-yà tsá Gulu yá*
 ASSC many child:PL beget-beget DEF Gulu DEM
 ‘Many children were begot by that Gulu.’

- (45b) *xiyá xiyá skwì txàf-í tà ná*
 corn corn thing expel-1SG PREP DEM
dêlèwèr ná ná skwì 3 nghá-nà-ghá-tsí
 notebook DEM DEM thing three look-DEM-D:GO-3SG
 ‘What I have written about in this notebook regards three things.’ (Frajzyngier with Shay 2002. The sentence is drawn from written sources with subsequent tonal notation added.)

If there is only one argument in the clause, the relevant distinction is between the controlling and the affected argument rather than between the subject and the object.

In the imperfective aspect, the noun phrase that precedes the verb is interpreted as the subject if the verb has the potential object marker (PO) *-ay*. The subject is interpreted as the non-controlling argument by the absolutive form of the verb (suffix *-kú*):

(46) *mbàd' ká kɾì kà dɔ́-áy*
 then COMP dog SEQ thresh-PO
 ‘And Dog kept threshing.’

mbàd' ká xìyà kà dɔ́-ákù-á-nì
 then COMP corn SEQ thresh-ABS-GEN-3SG
 ‘And corn kept threshing.’ (e.g. in a machine)

ndá sn-í ká tsá marriage-xà yá tà
 STAT hear-1SG COMP DEF marriage (Fr.)-PL DEM IMPF

mággá-kú gà xdí ká-xàn mándí bángál-xà yá
 make-ABS PREP Hdi COMP-3PL like marriage (Ful.)-PL DEM
 ‘I have heard that marriages are being made in Hdi.’

Mutsuvan

In Mutsuvan, if two arguments are present in a pragmatically neutral clause, the first argument is the subject; the second argument is the object. Thus the sole coding means in this kind of clause is linear order (all data from Eric Johnston, p.c., analyses and interpretation by this author):

(47) *mà ʔát hàdá ndé*
 PAST bite dog man
 ‘The dog bit the man.’

If the subject is fronted, the remaining object is marked by the preposition *tá*. This enables the interpretation of the fronted noun phrase as the subject:

(48) *à mí hàdá ʔát tá ndé*
 PRF dog bite OBJ man
 ‘The dog had bitten the man.’

ndé dà ʔát tá hàdá
 man PRES bite OBJ dog
 ‘The man is biting the dog.’

Thus the preposition in Mutsuvan is deployed only when the other coding means, viz. the linear order V NP NP, is not available. This fact is significant for two reasons: (1) it demonstrates the existence of the principle of functional transparency, and (2) it demonstrates that configuration is chosen first as the coding means, but when linear order is not available, a preposition is grammaticalized to code the grammatical relation.

The interpretation of the fronted noun phrase as the object is assured by the suffix *a*, probably the goal marker added to the verb (cf. Frajzyngier 2005, and by the fact that the noun phrase following the verb is interpreted as the subject:

(49) *hé pít-á nzè*
 corn harvest-GO 3INDEF
 ‘The corn is being harvested.’ (right now)

ándàva ná’w-á sónù
Cyperus esculentus L. (Cyperaceae) eat- GO Sunu
 ‘Sunu eats *Cyperus esculentus*.’

The evidence that the ending *a* is not the verbal ending before the subject or an inherent ending of the verb is provided by clauses in which the subject is not preceded by any vowel, as in example (47), repeated here for convenience:

- (50) *mà ʔət hàda ndé*
 PAST bite dog man
 ‘The dog bit the man.’

The absence of any preposition before the noun phrase following the verb is important because it indicates that this noun phrase is the subject.

The limited Mutsuvan data demonstrate the existence of the principle of functional transparency. Moreover, the coding means are deployed only when the principle of functional transparency may be violated.

Hona

In Hona, the fronted argument is coded as the subject through the third-person subject suffix added to the verb. Recall that when a nominal subject follows the verb, there is no third-person subject suffix added to the verb:

- (51) *Ali ʔe-n-də*
 Ali slaughter-OBJ-3SG
 ‘Ali slaughtered it.’

The fronted argument is coded as the object through object suffixes that follow the verb:

- (52) *mbe-nè iya-dĩ fwo-né-n*
 enjoy-1SG food-DEF cook-1SG.DAT-3SG.OBJ
 ‘I liked the food that she cooked for me.’

6.2.5 Summary of the coding with fronted arguments

Interpretation	Wandala	Hona	Hdi	Mutsuvan
Fronted NP as S	If the NP following the verb is coded as O	If the verb has subject marker	If the NP following the verb is coded as O	If the NP following the verb is coded as O
Fronted NP as O	If the NP following the verb is coded as S	If the verb has object marker	∅	If the NP following the verb is coded as S
Postverbal NP as S	Phrase-final form of the verb	∅	∅	Goal marker
Postverbal NP as O	bare stem (prf) -a (impf)	Subject pronoun on the verb	preposition <i>tá</i>	preposition <i>t</i>

7 The choice of formal means

When the need to code a given function arises, speakers have a large choice of means: linear order; phonological means (tone, stress, pauses, intonation, vowel deletion and retention, vowel harmonies); lexical categories and subcategories; and individual lexical items, such as specific nouns, specific verbs, or specific numerals and quantifiers, linear order, inflections, etc.

Linear order has an obvious limitation in that it allows the coding of two functions at best. This is the case with English NP NP constructions, where the functions of both NPs are coded, the first as head and the second as modifier. We can have either ‘a house tree’ or ‘a tree house’. Linear order is therefore quite limited in the numbers of distinctions it can code.

Another means of coding is through tonal changes. Here the number of distinctions coded depends on the number of underlying tones in the language. But even in a two-tone language the number of distinctions coded by tone can be twice two for any functional domain. Thus if low tone on the verb codes one tense, the high tone can code another tense. If a low tone pronoun codes indicative modality, the high tone can code deontic modality. Compare the examples from Hdi cited earlier in this study.

Similar constraints operate with respect to other prosodic means, such as stress and pauses, each of which offers a binary distinction at best.

Thus, when any of the binary distinctions is overlaid on some other coding means, e.g. segmental grammatical morphemes or word order, the number of coding means can double rather than being simply augmented by two for the language as whole.

For the case of the four Chadic languages with the verb in the clause-initial position, the position of the nominal argument before the verb in combination with other coding means codes topicalization or focus or backgrounding of the given noun. This position is therefore not available to code grammatical relations, such as subject or object. The formal means' having a function already assigned at the initial state plays a fundamental role in the choice of the formal means for the grammaticalization of other functions. Very simply, languages must look for other formal means.

In most languages, when a distinction for more than two relations between any two elements is called for, a means with richer coding possibilities is required. Drafting a lexical item to code the desired function is one of the ways in which languages enrich the coding means available.

8 Grammaticalization and language change

The implications of grammaticalization for language change follow from the grammaticalization of functions much more than from the grammaticalization of a given coding means. Once a function has been grammaticalized, the loss of the coding means that codes this function may have one of two consequences: (a) the given function disappears, or (b) the language grammaticalizes other means to code the same function. The loss of the coding means for a given function is paradoxically often a consequence

and a historical result of grammaticalization in that the form is used more frequently, and therefore it becomes predictable, allowing increased phonological reduction.

Grammaticalization theory should address the issue of which grammaticalized functions are allowed to disappear and which functions are provided with new means of coding.

One interpretation is that functional domains, once grammaticalized, remain all the time, as claimed by Aikhenvald and Dixon 2001. But is this true? Historical linguistics attests to functions that have disappeared from grammatical systems of unrelated languages. A well-known example is the disappearance of the category ‘dual’ from many Indo-European languages and of singular and plural distinction in the second person in English. It would appear that functions that are used extensively, e.g. functions that are coded in many types of utterance, are provided with new means of coding once the old means disappear. Functions with a limited use appear to be more vulnerable once their coding means disappear. We sorely lack cross-linguistic data to generalize on the retention and disappearance of functions even in languages with long written traditions. In languages without a written tradition, we even lack a reliable methodology to determine whether a given function is a retention or an innovation.

9 Grammaticalization as a research tool: Methodological considerations

Grammaticalization, as understood here, can also be used as a diagnostic tool in semantic analysis. The underlying principle is that if a function has been grammaticalized in a language, this function was not otherwise coded by lexical or grammatical forms of the language at the initial state, unless the motivation was the principle of indirect means.

With respect to semantic analysis, grammaticalization provides the evidence that lexical items that appear to have the same referent across languages have different semantic properties. The semantic properties are those that have been grammaticalized in a given language, either through the grammatical system or through the lexical categories and subcategories. Languages may grammaticalize different lexical categories as required by the types of predications they have. The grammaticalization of a given functional domain can be used as a diagnostic tool for semantic analysis, i.e. for the discovery of the meaning of a given lexical item in a given language. Thus the noun ‘room’ in English is

not inherently +LOCATIVE, in the sense of having this semantic feature, whereas its equivalent in Mina is.

10 Conclusions

Against the recent vigorous and extensive challenges to grammaticalization research and grammaticalization theory in its narrow sense of lexical item or construction to grammatical morpheme change (Campbell and Janda 2001, Campbell 2001, Newmeyer 2001, Joseph 2001, 2004), the present study proposes that grammaticalization, as here defined, is at the core of any linguistic theory that aims to explain why languages are similar and why they are different.

The study of grammaticalization can contribute to the foundations of linguistic theory by explaining the emergence of lexical categories, of word order as a coding means, of phonological means, in addition to the traditional scope, i.e. the emergence of grammatical morphemes from lexical items and constructions.

Within the expanded scope of grammaticalization, the motivation for each specific instance of grammaticalization becomes the most important object of research. Once the motivation is known, the next object of investigation is the choice of the means to code a given function.

Acknowledgments

Research on this study was supported by an award from the Charles and Jane Butcher Foundation and by a National Science Foundation grant Nr. 0439940 for the study of Wandala and Giziga. I am grateful to Eric Johnston for sharing his data on Mutsuvan. Erin Shay read an earlier version of this paper and provided valuable comments. Henry Tourneux provided the correct identification for the plant *Cyperus esculentus*, which in Cameroon is often referred to by the French name *chiendent* --a term that in France designates a different plant. My thanks to Maria Jose Lopez-Couso, whose sharp eye spared me an embarrassing error. I am grateful to Marian Safran for editorial work. I am grateful to an anonymous referee whose close and critical reading and constructive comments prompted me to revise the paper, tighten up the presentation, and check

additional references. This study would not have been possible without the speakers of Wandala, Mina, Hdi, Giziga, Gidar, Hona, and Kafa who over the years were willing to sit with me and help me in the study of their languages. Any errors of fact or interpretation remain my sole responsibility.

References

- Aikhenvald, Alexandra Y. 2000. *Classifiers*. Oxford: Oxford University Press.
- Aikhenvald, Alexandra Y., and R.M.W. Dixon (eds.) 2001. *Areal diffusion and genetic inheritance: Problems in comparative linguistics*. Oxford; New York: Oxford University Press.
- Blevins, Juliette. 2005. *Evolutionary Phonology: The Emergence of Sound Patterns*. Cambridge: Cambridge University Press.
- Campbell, Lyle. 2001. "What's wrong with grammaticalization?" *Language Sciences* 23. 113-161.
- Campbell, Lyle. 2004. *Historical Linguistics*. 2nd ed. Cambridge, Mass.: MIT Press.
- Campbell, Lyle, and Richard Janda. 2001. "Introduction: Conceptions of grammaticalization and their problems." *Language Sciences* 23.93-112.
- Croft, William. 2001. *Radical construction grammar: Syntactic representation in typological perspective*. Oxford: Oxford University Press.
- Dixon, R.M.W. 1977. "Where have all the adjectives gone?" *Studies in Language* 1. 19-80.
- Fischer, Olga, Muriel Norde, and Harry Perridon. 2004. *Up and down the cline: The nature of grammaticalization*. Amsterdam/Philadelphia: Benjamins.
- Frajzyngier, Zygmunt. 1987. "Grammaticization through analysis." In *Papers from the 2nd Pacific Linguistics Conference*. Ed. Scott DeLancey and Russell Tomlin, 125-140.
- Frajzyngier, Zygmunt. 1991. "The de dicto domain in language." In *Approaches to Grammaticalization*. Ed. Elizabeth C. Traugott and Bernd Heine, 1:219-251. Amsterdam/Philadelphia: Benjamins.

- Frajzyngier, Zygmunt. 1996. *Grammaticalization of the Complex Sentence: A case study in Chadic*. Complementary Series to the Study in Language. Amsterdam/Philadelphia: Benjamins.
- Frajzyngier, Zygmunt. 1997. "Pronouns and agreement: Systems interaction in the coding of reference." In *Atomism and binding*. Ed. Hans Benis, Pierre Pica, and Johan Rooryck, 115-140, Dordrecht: Foris.
- Frajzyngier, Zygmunt. 2004a. "Principle of functional transparency in language structure and language change." In *Linguistic diversity and language theories*. Ed. Zygmunt Frajzyngier, David Rood, and Adam Hodges. Amsterdam/Philadelphia: Benjamins.
- Frajzyngier, Zygmunt. 2004b. "On segment counting, OCP, and heavy segments." *Folia Orientalia* 40.93-134.
- Frajzyngier, Zygmunt. 2005. "La catégorie augmental 'but' en tchadique." In *Les Langues Chamito-Sémitiques*. Ed. Amina Mettouchi and Antoine Lonnet. *Faits de Langues*. Paris:Ophrys.
- Frajzyngier, Zygmunt. (Forthcoming). *A Grammar of Gidar*. Frankfurt: Peter Lang.
- Frajzyngier, Zygmunt, and Bill Jirsa. 2005. "The principle of indirect means in language use and language structure." *Journal of Pragmatics* 38.4. 513-542.
- Frajzyngier, Zygmunt, Eric Johnston, with Adrian Edwards. 2005. *A Grammar of Mina*. Berlin/New York: Mouton de Gruyter.
- Frajzyngier, Zygmunt, Holly Krech, and Armik Mirzayan. 2002. "Motivation for copula in equational clauses." *Linguistic Typology* 6.2.155–198.
- Frajzyngier, Zygmunt, and Jan Mycielski. 1998. "On some fundamental problems of mathematical linguistics." In *Mathematical and computational analysis of natural language*. Ed. Carlos Martin-Vide, 295-310. Amsterdam/Philadelphia: Benjamins.
- Frajzyngier Zygmunt, with Erin Shay. 2002. *A Grammar of Hdi*. Berlin: Mouton de Gruyter.
- Frajzyngier Zygmunt, and Erin Shay. 2003. *Explaining language structure through systems interaction*. Amsterdam/Philadelphia: Benjamins.
- Giacalone-Ramat, Anna, and Paul Hopper (eds.) 1998. "Introduction." In *The Limits of Grammaticalization*, Typological Studies in Language 37. Amsterdam/Philadelphia: Benjamins.

- Givón, Talmy. 1990. *Syntax. A functional-typological introduction*. Vol. 2. Amsterdam: Benjamins.
- Goldenberg, Gideon. 1971. "Tautological infinitive." *Israel Oriental Studies* 1.36-85.
- Greenberg, Joseph H. 1978. "How does a language acquire gender markers?" In *Universals of human language*. Ed. Joseph H. Greenberg, Charles A. Ferguson, and Edith Moravcsik. 4:47-82. Stanford: Stanford University Press.
- Haiman, John. 1994. "Ritualization and the development of language." In *Perspectives on grammaticalization*. Ed. William Pagliuca, 8-28. Amsterdam & Philadelphia: Benjamins.
- Haspelmath, Martin. 2004. "On directionality in language change with particular reference to grammaticalization." In Fischer et al. 2004.
- Heine, Bernd, and Tania Kuteva. 2002. *World lexicon of grammaticalization*. Cambridge: Cambridge University Press.
- Heine, Bernd, Ulrike Claudi, and Friederike Hünemeyer. 1991. *Grammaticalization. A conceptual framework*. Chicago: University of Chicago Press.
- Heller, Daphna. n.d. Possession as a lexical relation: Evidence from the Hebrew construct state. atar.msc.huji.ac.il/~english/IATL/18/Heller
- Hopper, Paul J. 1987. "Emergent grammar." *Berkeley Linguistic Society* 13.139-157.
- Hopper, Paul J., and Elizabeth Closs Traugott. 1993. *Grammaticalization*. Cambridge: Cambridge University Press.
- Joseph, Brian D. 2001. "Is there such a thing as "grammaticalization"?" *Language Sciences* 23.163-186.
- Joseph, Brian D. 2004. "Rescuing traditional (historical) linguistics from grammaticalization theory." In Fischer et al. 2004.
- Lehmann, Christian. 2002 (1982). *Thoughts on grammaticalization: A programmatic sketch*. Erfurt: Seminar für Sprachwissenschaft der Universität (ASSidUE, 9). 2nd ed., rev.
- Matthews, Peter. 1997. *The Concise Dictionary of Linguistics*. Oxford: Oxford University Press.
- Newmeyer, Frederick J. 2001. "Deconstructing grammaticalization." *Language Sciences* 23.187-229.

- Rijkhoff, Jan. 2002. *The noun phrase*. Oxford: Oxford University Press.
- Roberts, Ian, and Anna Roussou. 2003. *Syntactic change: A minimalist approach to grammaticalization*. Cambridge: Cambridge University Press.
- Samuels, M.L. 1972. *Linguistic evolution, with special reference to English*. Cambridge: Cambridge University Press.
- Shay, Erin. *Grammar of Giziga*, in progress.
- Shibatani, Masayoshi. 1991. "Grammaticalization of topic into subject. In *Approaches to Grammaticalization*." Ed. Elizabeth C. Traugott and Bernd Heine, 2:93-134. Amsterdam/Philadelphia: Benjamins.

Abbreviations

1	First person
2	Second person
3	Third person
ABS	Absolutive
AFF	Affected argument
ALL	Allative
ANAPH	Anaphor
ASSC	Associative
BCKG	Backgrounding
COMP	Complementizer
D	Dependent (aspect)
DAT	Dative
DEF	Definite
DEM	Demonstrative
DET	Determiner
EE	End of episode
GEN	Marker of modification
GO	Goal orientation

HYP	Hypothetical
IMPF	IMPERFECTIVE
IMPR	Imperative
IN	Extension coding the movement in
INCL	Inclusive
INDEF	Indefinite
INF	Infinitive
INN	Inner space extension
INTENS	Intensifier
IP	Independent pronoun
IRR	Irrealis
M	Masculine
NEG	Negative
NORM	Normative
OBJ	Object marker
ON	Extension coding location 'ON'
OUT	Extension coding the movement out
PART	Partitive
PAST	Past tense
PL	Plural
PO	Potential object
POS	Point-of-view of subject
PRED	Predicator
PREP	Preposition
PRES	Present tense
PRF	Perfective
PRO	Pronoun
PROG	Progressive
PVG	Point of view of goal
Q	Question
REF	referential

REL	Relative marker
SG	Singular
SEQ	Sequential
STAT	Stative
SUBJ	Subjunctive
TOP	Topic marker
VENT	Ventive

¹ An anonymous reader asks whether ‘this means that every pattern of language structure fits under one or the other of these two notions’? At present I would say yes, bearing in mind, however, that there are languages where the notion of a lexical item may be difficult to establish, and where ‘lexicalization’ includes formation of complex structures from the available syntactic means, and. Such is the case in some Amerindian languages. The instance of ritualization and conventionalization in language as described in Haiman 1994 (pointed out by the anonymous reader) may also lie outside the lexicalization and grammaticalization dichotomy.

² An anonymous reader points out that existence of synonymous constructions contradicts this statement. But do we really have synonymous constructions in languages, i.e. constructions that in addition to any referential meaning have also the same function within a clause, a sentence, and in discourse, and used by the same speakers? Any claims of such synonymy should be proved rather than asserted.