Towards a non-aprioristic syntactic theory

Zygmunt Frajzyngier

Department of Linguistics, University of Colorado
Boulder, Colorado 80309

1 Aims, the scope, and assumptions of the study

The first goal of this study is to outline the basic components of a non-aprioristic, empirically based syntactic theory aimed at explaining similarities and differences among languages. The second aim is to provide evidence for the existence of these components. The third is to demonstrate the explanatory possibilities of the theory and its constraints, i.e., what kinds of facts it is able to explain and what kinds of facts it is unable to explain at present. The proposed approach contributes to the explanation of differences among languages.

The study is a conceptual and an empirical extension of hypotheses and argumentations proposed in Frajzyngier and Shay 2003. Within the proposed theory, language is viewed as a means to create meaning. This contrasts with approaches in which language is viewed as a means of connecting meaning with phonetic realization, e.g.: "Language is an obvious interface of (at least) sound and meaning, and LF [Logical Form] representations code whatever syntax inputs meaning [sic]" (Lasnik et al. 2005: 216. For a similar view, viz. that language connects sound and meaning, see Jackendoff 2002: 282. The difference in approach is quite important in that, within the proposed model, if the grammatical system does not encode a certain meaning, the speakers of the language do not have to attend to it, as already noted in Jakobson's well-known statement "Languages differ essentially in what they must convey and not in what they may convey" (Jakobson 1959). A simple illustration of the difference: Many Chadic languages encode the categories 'ventive', roughly indicating movement toward the place of speech or some other deictic center, and 'andative', roughly indicating movement away from the deictic center. The two categories are coded on verbs of movement and also on other verbs. Consequently, each utterance in a language that encodes these categories has to indicate whether the event involved movement of the subject or object to place of speech or some other place. One can express such distinctions in English and other Indo-European languages, but the listener does not expect such information unless it has been requested. A periphrastic coding of such a distinction in every sentence would be considered an affectation at best, but more likely an instance of speech pathology requiring a correction.

Having stated that grammar creates meaning, we need to explain why languages are similar and why they are different. Some similarities across languages can be explained by common human neurological make up, by characteristics of the task of communication, or by similar needs of human interactions. These common human characteristics are too general to explain any specific grammatical similarity, such as the existence of similar grammatical categories across languages or, indeed, the very existence of grammar. The common needs and characteristics of human communication

cannot explain the differences among languages. Frajzyngier and Shay 2003 and Frajzyngier 2004 postulate the existence of a principle of functional transparency, a fundamental characteristic of natural languages, that accounts for some design similarities across languages. The present study provides new evidence for the existence of this principle and demonstrates its explanatory power. The principle allows one to explain facts for which other contemporary and traditional theories do not provide an explanation.

Neither the principle of functional transparency nor the common cognitive abilities of humans can explain differences across languages. Within the theory proposed in Frajzyngier and Shay 2003 and developed in the present study, two factors are partly, though not exclusively, responsible for differences among languages: (1) the fact that languages code different functional domains and/or sub-domains, and (2) the fact that languages have different coding means available.

One of the important outcomes of the proposed theory is that many differences across languages can be explained by a relatively small number of language-internal properties. This study shows how the presence of certain properties in selected languages can be explained in terms of language-internal properties and the principle of functional transparency. The study is organized as follows: The next section (2) consists of a brief review of how selected linguistic theories address the issue of similarities and differences across languages. This is followed by the formulation of, and evidence for, the principle of functional transparency (Section 3). Section 4 is a discussion of selected formal means that have not been discussed in the literature or that require additional explanation. Section 5 describes how the various formal means interact in the coding of functional domains. Section 6 describes implications of the proposed framework and proposes questions for future research.

2 How various theories deal with similarities and difference across languages

This brief review of current linguistic theories addresses just one issue: how the theory attempts (if at all) to explain similarities and differences across languages. It is not an attempt to present the totality of the theories, of their assumptions, of their possibilities, or of their constraints. It is also not an evaluation of how well a theory explains facts of a single language. The formal apparatus proposed by various theories is also outside of the scope of this review. The discussion below provides, however, a discussion of some of the implicit or explicit assumptions of various theories that prevent them from explaining similarities and differences among languages.

Within the Minimalist program, similarities among languages are presumed to be manifestations of universal grammar. Differences among languages are supposed to be accounted for by choosing different values for the same parameters, as proposed in principle and parameters theory (Chomsky and Lasnik 1991, Culicover 1997). One of these parameters is directionality in the formation of head-dependent structures. In this view, some languages are viewed as left-branching and other languages as right-branching. Behind this parameter is an assumption that all elements of an utterance are either the heads or the dependents of some larger constructions. In the present work (section 3.1), this assumption is shown to be incorrect with respect to one parameter that is of primary importance in the Minimalist Program's 'merge operation' and is shared by

the head-marking versus dependent-marking approach to language structure, as proposed in Nichols 1996, and Nichols and Bickel 2006). The principles and parameters theory does not explain why some languages choose one value along a given parameter and other languages choose another value. If such parameters do exist, one would still need to explain why different languages make different choices along those parameters. The present paper provides a means of explaining such choices.

Neither usage-based models nor construction grammar address the fundamental issues of similarities or differences among languages. The theoretical assumptions of the usage-based models (Langacker 2000, Barlow and Kemmer 2000, Tomasello 2003, Bybee 2009) do demonstrate that frequency of use leads to grammaticalization, giving support to long-held assumptions about the role of frequency of use. But the usage-based theories do not explain why different languages have different frequencies of use. More fundamentally, usage-based theories are quite vague with respect to the question of what it actually is that the speakers use. If the form of a language were actually determined by the usage, all languages would in principle be very similar unless drastic differences in language use could be demonstrated across cultures. No such differences have been demonstrated. In its present shape, the usage-based models would actually preclude any kind of principled explanation of differences among languages.

The construction grammar approach explicitly aims at explanation of constructions within a single language. Similarities among languages are presumably explained by the existence of similar constructions across languages (Goldberg 2006). Construction grammar does not explain why languages have some constructions rather than others.

Dependency grammar, as presented in Tesnière 1966, does not address the issue of similarities and differences across languages.

Lexical-functional grammar, which postulates some general principles that presumably are shared by all languages, does not explain specific similarities or differences across languages (Falk 2001). There appears to be an underlying assumption within lexical functional grammar that languages code the same functions, and that differences among languages consist of different formal means deployed to code the same functions.

Given the paucity of theoretical explanations for differences across languages, there is a need for an empirically based theoretical model to account for some of the differences and some of the similarities across languages. What follows is a proposal for such a model.

3 An outline of the theoretical framework

3.1 The notion of syntax and semantics

The generative tradition, from its very beginning more than fifty years ago to its most modern incarnation, has retained the traditional conception of syntax as a means of forming grammatically correct sentence out of the lexicon. Within the minimalist program, this is the gist of the first operation in syntax, namely ‘merge’ (Chomsky 1995). The same approach to syntax, as a means of forming sentences out of words, is taken in lexical-functional grammar (Falk 2001: 4). The fundamental problem with this approach is that the very object that those theories claim to describe, viz. the formation of sentences out of words does not seem to exist. There is no evidence of any kind that

speakers select some lexical items and then form sentences out of them. In the present study, I provide the empirical evidence that such an operation does not exist. Behind the traditional approach to syntax, as exemplified by generative grammar and lexical-functional grammar, lies another tacit assumption that much (all?) of the meaning in language is contained in and conveyed by lexical items. This assumption, which might have been justified for the state of knowledge in ancient grammatical tradition, cannot be retained today when we know that all other formal means existing in the language contribute in equal measure to the coding of meaning. In what follows, I present an alternative to the traditional approach to syntax and semantics.

3.2 Predication – a basic unit of semantic structure

One of the basic components of semantic structure is an entity that I shall call "predication".[i] It is a form or a set of forms that code one specific function. Predication is characterized by strict formal properties that may vary within the language depending on the properties of lexical items selected for the predication.

The realization of various predications follows the principle of systems interaction, whereby the distinguishing characteristic of a single predication is not coded twice within the same construction. Several linguistic forms may be called upon to encode one semantic predication. A predication may be realized by only one construction or by several constructions that are mutually complementary, representing the interaction of various coding means available within the given language (Frajzyngier and Shay 2003).

For example, adverbial predication in English may consist of a single lexical item if the lexical item is an inherent adverb, or it consists of a preposition and a noun if the lexical item selected is not an inherent adverb. Here are examples of adverbs of manner (all examples from English come from the London-Lund corpus; formal means indicating prosodic characteristics of speech are suppressed):
Inherent adverb 'fast':
10_4_3 <901 a> and Archie coming now **fast** again on the stand side.

Derived adverb 'motionless' from the noun 'motion':
10_5_0 <975 b> they stand **motionless**

Inherent adverb of intransitive verbs: 'very smart' :
10_5_0 <976 b> looking very smart

Adverbial phrase marked by the preposition 'with' because 'strict discipline' is a noun phrase:
10_5_0 <977 b> **with strict discipline**

A predication is different from a proposition in that a predication may have different modalities. Moreover, predication is not subject to truth-theoretic analysis. An individual utterance, clause, sentence, or whatever formal entities a given language has, may be composed of one or several predications, each coding a different facet of the meaning of the utterance. The notion of the predication structure of an utterance may be somewhat reminiscent of the generative semantics ideal of representing the meaning of a sentence. Unlike the generative semantics approach, does not take for granted the existence of any meaning, and postulates that every meaning coded in a given language has to be discovered.

The importance of predication as a unit of semantic structure is that it allows us to explain the wide variety of language structures. It also allows us to discover the semantic features of individual lexical items and grammatical morphemes. A predication is not the same as a construction, as a single predication may be realized by a variety of constructions, depending on the properties of the lexical items used. The following discussion of locative predication in Mina (Central Chadic) has four purposes:

(1) to provide the evidence for the existence of the category predication.

(2) to demonstrate that the speaker chooses a variety of formal means available in a given language to code the predication he/she wants to express. The formal meansincludes the choice of lexical items that the speakers wants to talk about.

(3) to demonstrate that the choice of some lexical items depends on the type of predication involved and on the choice of other lexical items, a possibility not envisaged by any of the existing linguistic theories.

(4) to demonstrate that the properties of lexical items within a language are not arbitrary but rather depend on the type of predications coded in the language. This runs contrary to the usual assumption that properties of individual lexical items are arbitrary, save for the possibility that some properties of lexical items reflect the speakers' cognitive systems. The outcome of this discussion is a demonstration that the operation 'merge' in the minimalist program, and the standard approach in lexical-functional grammar of combining words into larger elements, does not exist.

Locative predication in Mina codes movement to or from a place or the presence of an entity or an event at a place. The existence of locative predication as a separate predication in the language is proved by the strict interaction of the formal means in the coding of the predication. The formal means involved are properties of the predicates and nouns serving as locative complements, the predicator *á*, and the preposition *n*.

In locative predications in Mina, if the predicate is not inherently locative, the locative predicator *a* must be used (all Mina examples are from Frajzyngier et al. 2005):

(1) a.	*i*	*ŋ*	*kə*	*ŋdəv-a*	***a***	*kayak*
	3PL	PREP	INF	fall-GO	PRED	earth

'They will fall down on the ground.' (written sources, hence no tones)

If the noun is not inherently locative, the locative preposition must precede the locative complement:

(1) b.	*tsáy*	*mə*	*tí*	*tí*	*nd-á*	*nástə́*	***nə̀***	*yəm*
	then	REL	look	look	go-GO	enter (F.)	PREP	water

'Then the one who was good at looking entered into water.'

If neither the predicate nor the complement is inherently locative, both the locative predicator *á* and the locative preposition *n* are deployed (the preposition n may be followed by the epenthetic vowel if the phonological environment so requires):

(1) c.	*ván*	*ɗá*	*rà*	*mənə́*	***á***	***nə̀***	*lùmò*
	rain	fetch:GO	D.HAB	like	PRED	PREP	market

'It was raining from the direction of the market.'

If both the predicate and the complement are inherently locative, no predicator or preposition is deployed:

(1) d. *ábə̀ nd-á ngə̀n wùtá*
ASSC go-GO 3SG village
'Then she returned to her village.'

There are a number of implications of the analysis of locative predication as proposed above. The first one concerns the role of syntax, understood as the totality of grammatical means. The role of syntax is not to put words together, as postulated in traditional grammar and in the contemporary generative grammar and lexical-functional grammar, but rather to encode a predication. Syntax selects among linear order, inflectional morphology, lexical items in grammatical function and other grammatical means to interact with the lexical items chosen in order to code the desired predication.

The types of syntactic means deployed, viz. the predicator *á* and the preposition *n* in the clause depend on the type of predication coded in the language, on the properties of this predication, and on the semantic features of the lexical items involved in the predication.

Properties of lexical items within a given language reflect the types of predications grammaticalized in the language. Contrary to the traditional approach perpetuated in contemporary linguistic theories, the properties of lexical items are not arbitrary but rather are imposed by the types of predications existing in the language. This explains why lexical items that have the same reference across languages or across cognitive systems may have different semantic features. Thus, in Mina, some nouns and verbs are [+locative] and the others are [-locative]. Their value is not determined by what they refer to in reality, or by the cognitive systems as favored by proponents of 'cognitive grammar' but rather by their role within the locative predication. It appears that there are neither inherently locative predicates nor inherently locative complements in English because locative predication does not exist as a separate type of predication that is different from all other types of predications. Locative meanings illustrated in Mina are subsumed in English under the functions coded by prepositions such as 'to', 'from', 'at', 'in'. Though some nouns, e.g. 'home', can occasionally be used without a preposition, the evidence that such a noun is not inherently locative is that it is preceded by a preposition when functioning as a locative complement.'

2_13_0 <952 A> they got three dogs at the home

The prepositions 'from' and 'at' appear to have only a locative meaning. Other prepositions can mark a variety of complements, some of which may be locative. Thus, the preposition 'to' marks infinitive verbs, a function unrelated to locative predication in any sense:

12_7_0 <832 a> it is hard **to** imagine a more exhaustive report

12_7_0 <838 a> and their representation in public
12_7_0 <839 a> are deemed **to** be both immoral and offensive.

The preposition 'at' can mark non-locative complements:

12_7_0 <654 a> may be led to believe that the authorities have lent their approval
12_7_0 <655 a> or **at least** acquiescence

The verb ‘go’ and its tensed alternative ‘went’, whose equivalent in Mina is inherently directional and hence locative, can be used in English with non-locative complements:

11_5_0 <810 f> and I think a lot of people in Scotland
11_5_0 <811 f> **went** along with that

12_1_2 <399 a> so that the cheap things looked were highly priced
12_1_2 <400 a> and the valuable things **went** for a song

12_2_2 <1312 a> his father **went** virtually bankrupt (London-Lund corpus)

Predications in individual languages may be members of various functional domains, as described in the next section.

3.2 Notion of functional domains

A fundamental component of the non-aprioristic linguistic theory is the notion of functional domain defined as follows:

(a) A domain D is a class of mutually exclusive types of expressions constructed by a set of specific means of coding M(D).
(b) The types of expressions within the domain D have a specific pragmatic or semantic function in common.
(c) A domain D is a class of types of expressions with a certain set of meanings D* such that D* is disjoint with E* for any domain E ≠ D of the same language (Frajzyngier and Mycielski 1998)

No two members of the same domain can co-occur in the same predication. For example, if a language codes functions in the domain of tense, one cannot have a future and a past tense marker within one simple clause. In a language that codes the perfective and the imperfective aspect, one cannot code the perfective and the imperfective aspect within one simple clause. In a language that codes nominal number, one cannot have a singulative and a plural marker added to the same noun. Co-occurrence restrictions of this type are the main criterion for determining whether two or more forms belong to the same or to different domains.

Descriptive grammars of individual languages are seldom explicit with respect to what functional domains are coded in the language and what is the internal structure of those domains. Nevertheless, they do allow us to detect that each grammatical system codes a finite number of functional domains. Thus, some languages code aspectual systems, other languages code tense systems, some languages code both aspectual and tense systems, and some languages may have only one system of coding both temporal aspectual distinctions. While the number of functional domains coded in the language at any one point is finite, this does not mean that it is a closed set. Some functional domains may disappear and others may emerge. Thus, an aspectual system may evolve into a tense and aspectual system, or it may disappear altogether, leaving only a tense system in

place. The emergence of the system of definite and indefinite articles in Germanic and Romance languages is an example of the emergence of a sub-domain within the domain of reference.

Given the cultural needs of human communication and the requirements of discourse, we may expect some functional domains to be encoded across languages more often than others. For example, modality, a domain that may include means of conveying the message as true, asking questions about the truth, asking content questions, and giving commands appears to be attested in all languages described so far. Languages might also be expected to encode certain kinds of relationships within a proposition, but here a great deal of variation is attested. Some languages code semantic relations between the predicate and arguments (Mithun 1991), while other languages distinguish between the grammatical relations of the two arguments and leave the semantic relations unspecified. This is the case in languages that have the categories subject and object. Some languages have a mixed system in which both semantic relations and grammatical relations are coded at the same time (Frajzyngier and Shay 2003, Frajzyngier and Munkaila 2004).

Even if languages code the same functional domains, they may differ with respect to the internal structures of the domains. Consider content questions about human participants. In English, in the domain *de dicto* (Frajzyngier 1991), such questions do not distinguish between singular, plural, male, or female participants:

1_2_2 <1007 B> so **who** will it go to

The form 'who' allows for the possibility of an answer involving plural participants as evidenced by the following exchange:

1_3_0 <876 c> well, **who**'s in the senior common-room

1_3_0 <877 A> all the academic staff (London-Lund Corpus)

There are periphrastic means in English to encode presuppositions about participants in questions in the domain *de re*:

(2) a. '**Who** was the man/the woman who . . .'
(2) b. '**Who** were the people who . . .'

In Gidar (Central Chadic), a content-question phrase consists of the form coding person and a copula, which distinguishes between the masculine, feminine, and plural forms. Consequently, content questions may encode the gender and number of the participants. If the speaker has no presuppositions about the gender and number of the participants, the masculine form of the question phrase is used:

(3) a.	*náwá-**y***	*də-də*	*sə́-n*	*wáy*	*dì*
	who-COP.M	3M-cook	DAT-3M	food	SQ
	'who (m) cooked food for him?'				

(3) b. *náwá-**t*** *tə̀-də̀* *sə́-n* *wáy* *dì*
who-COP.F 3F-cook DAT-3M food SQ
'who (f) cooked food for him?'

(3) c. *mày* *náwá-**ŋ*** *də̀-dà* *zzúw-àŋ* *sə́* *jáaɓè* *dì*
ASSC.PL who-COP.PL 3M-D.PROG come-PL from Djabe SQ
'who are the people coming from Djabe?'

Similarly, in questions about the dative argument, the marker of the dative is followed by the gender marker indicating the speaker's presupposition about the identity of the dative argument:

(4) a. *sə́-**t*** *náwá* *kə̀-psə́-t* *bàrdáw* *ván* *dì*
DAT-3F who 2SG-give-3F hoe DEF SQ
'to whom (f) did you give the hoe?'

(4) b. *sí-**t*** *mày* *náwá* *kə̀-psí-t* *bàrdáw* *ván* *dì*
DAT-3PL ASSC.PL who 2SG-give-3PL hoe DEF SQ
'to whom (pl) did you give the hoe?' (all examples from Frajzyngier 2008)
(The underlying form of the third-person plural marker is *tì*. The presence of the high front vowel causes fronting of the preceding vowel.)

The fact that languages may code different meanings in their grammatical systems is one of the factors responsible for the differences among languages. What meanings are coded by the grammatical system cannot be predicted in the current state of linguistic knowledge. We have no hierarchies concerning which meanings are encoded more often and which are encoded less often. Communicative needs may be responsible for the emergence of functional domains. The emergence of the sub-domains may be a by-product of the availability and the properties of coding means in a given languages, as is the case with the questions about human participants in Gidar. The emergence of individual constructions coding narrow meanings is often an opportunistic exploitation of the formal niches available in the language. An example is the emergence of the detrimental function of the English preposition 'on' in clauses of the type 'the car did not start on me'. The preposition has long been available to code the locative function combined with a specific spatial orientation. Its use with obviously non-locative complements, most often [+ human], led to the emergence of the malefactive function of the preposition.

Despite the significant accumulation of data and now also the convenient availability of data for linguistic typology (see Haspelmath et al. 2005), a typology of functional domains and their internal structures has yet to be produced. A necessary prerequisite for such a typology is a non-aprioristic analysis of the meanings actually encoded in individual languages.

3.3 Principle of functional transparency

The principle of functional transparency is postulated to be the fundamental principle in syntactic organization. Given its importance, the present study provides cross-linguistic

evidence for its existence and describes the explanatory potential of this principle. The principle of functional transparency is grounded in the communicative function of language. It is testable on a language of any type and is therefore falsifiable.

The form of an utterance is a function of the types of predications (functions) grammaticalized in the language and the lexical and grammatical means available for the coding of these predications. The form of the utterance obeys the principle of functional transparency:

> Within a given functional domain, the role of all elements must be transparent to the hearer. This transparency relates only to the functions that have been grammaticalized in a given language rather than to the hearer's need to understand the relationships in reality outside of language. The functional transparency is assured by the formal means existing in a given language, which may include the following (as listed in (Frajzyngier and Shay 2003):
>
> Phonological means, including prosodic characteristics (tone, intonation, pauses and their absence), segment reduction
> Lexical categories and subcategories
> Inflectional means on all lexical categories
> Serial verb constructions
> Prepositions and postpositions
> Deployment of lexical items to code grammatical functions
> Linear order
>
> This list of coding means is considerably expanded in section 5 of the present study.

The evidence for the existence of the principle of functional transparency is provided by the fact that if functional transparency is violated, utterances are ungrammatical. This principle holds across languages. Here is the evidence and illustration of the operation of the principle of functional transparency with respect to roles of noun phrases in three languages: English, Polish, and Mina (Central Chadic).

3.3.1 English prepositions

In most studies of English, the existence of prepositional phrases is taken as given. The importance of the principle of functional transparency rests in the fact that it can explain why English has prepositional phrases. The analysis of English prepositions proposed in Frajzyngier and Shay 2003 is replicated here, since it is easier to illustrate an issue on a familiar language rather than on a language that is not familiar. The question to be answered is: Why are the subject and object in English not marked by prepositions while other noun phrases are marked by prepositions? The answer is that the functions of subject and object are marked by the positions before and after the verb, respectively. This preempts the use of these positions for the coding of other roles, with the exception of benefactive role, and the only other means available in English are prepositions (all data are from the London-Lund corpus). The role of position in coding the benefactive function is described in section 5.4:

(5) a. **in Indonesia** I had what I thought was chicken

(5) b. and he recommended they have some dish **on an Italian ferry**

Compare the ungrammatical utterances without prepositions:

(6) a. ***Indonesia** I had what I thought was chicken

(6) b. *and he recommended they have some dish **an Italian ferry**

Consider also temporal phrases. If the lexical item is an inherent adverb (whether simple or derived), it does not require a preposition:

(7) oh a friend of mine was saying **yesterday** that that

banks were telephoning the Bank of England **hourly**

If the lexical item is not an inherent adverb, it requires a preposition to indicate its role in the utterance:

(8) 1_3_0 <496 A> me and the Edinburgh girl
1_3_0 <497 A> got together **after dinner**
1_3_0 <498 A> late **in the evening**

Compare the use of the noun 'evening' as an argument rather than adjunct:

(9) 1_5_0 <51 B> and then you'd have **a whole evening** battling on

Hence, prepositions in English are deployed when other means of coding the role of the noun phrase are not available. The principle of functional transparency precludes the existence of systemic ambiguity in the language, i.e. ambiguity that would result from the coding of different functional domains through similar or identical means.

3.3.2 Polish genitive

The genitive case in Polish is deployed in a large number of predications. In this study I shall look only at the interaction of two of these: the modifying relationship among nouns and the obligatory marking of the second argument in negative predication. Negative predication may contain an overt negative marker or inherently negative verb. First, examples of the genitive construction in which the modifier has genitive case marking (all examples from sources to the Polish frequentative dictionary by Kurcz et al. 1990):

(10)

jeżeli	*ja*	*wypiję*	***herbatę***	***Ryszarda***
if	1SG	drink:1SG:FUT	tea:ACC	Richard:GEN

'If I drink up Richard's tea?'

druzgocąca ***przewaga*** ***wroga*** *to* *podstawowy* *warunek*
overwhelming:NOM advantage:NOM enemy:GEN DEM fundament condition
'The overwhelming advantage of the enemy is the fundamental condition.'

wyrazy ***czci*** ***i*** ***poważania***
expressions:NOM honor:GEN CONJ respect:GEN
'expressions of honor and respect'

Examples of genitive marking on the second argument ('object') of the negative predication:

(11)
prawa *nie* *miałeś*
right:GEN NEG have:2SG:M:PAST
'You did not have the right.'

po *prostu* *nie* *znoszę* ***widoku*** *samotnej* *filiżanki*.
PREP simply NEG support:1SG:PRES view:GEN solitary:GEN cup:GEN
'Simply, I cannot stand the view of a solitary cup.'

The single argument of the negative existential predication is also coded by the genitive case:

(12)
niczego *nie* *będzie,* *nie* *będzie* ***jublu***
nothing:GEN NEG be:FUT:3SG NEG be:FUT:3SG celebration:GEN
'There won't be anything, there will be no celebration.'

A negative predication combined with a genitival construction presents an interesting challenge to the principle of functional transparency in that there are two nouns marked by the genitive case. Given that word order in Polish does not code grammatical relations or the relations between noun phrases (case markers perform those functions), either noun could be the second argument of the negative predication or the modifier of a head. Hence, there is a potential for systemic ambiguity. The evidence for the existence of the principle of functional transparency is provided by the fact that whenever a systemic ambiguity obtains, such as in constructions with two genitives, the language deploys means to prevent systemic ambiguity. In Polish, these means involve the use of the position after the head noun to code the modifying function only. Here is the proof: Consider the following fragment from a poem:

(13)
Nie *znajdziecie* *na* *żadnej* *z* *map*
NEG find:FUT:2PL on none from map:PL:GEN

Tajemniczej ***ojczyzny*** ***poetów***
mysterious:SG:GEN fatherland:SG:GEN poet:PL:GEN
'On no map will you find the mysterious fatherland of poets.' (Tuwim, *Do generałów*, 1975 edition, p. 288)

The sequence of two noun phrases in the genitive case, *tajemniczej ojczyzny poetów* 'the mysterious fatherland of poets', cannot be separated by any other element present in the clause while retaining the same meaning. If a noun phrase in the genitive case is preceded by another noun, the first noun is the head of a noun phrase whose modifier is the noun in the genitive case:

(14)
Nie znajdziecie tajemniczej ojczyzny
NEG find:FUT:2PL mysterious:SG:GEN fatherland:SING:GEN
na żadnej z ***map*** ***poetów***
on none from map:PL:GEN poet:PL:GEN
'You will not find the mysterious fatherland on any of the maps of poets.'

Nie znajdziecie poetów na żadnej z ***map***
NEG find:FUT:2PL poet:PL:GEN on none from map:PL:GEN
tajemniczej ***ojczyzny***
mysterious:SG:GEN fatherland:SG:GEN
'You will not find poets on any of the maps of the mysterious fatherland.'

The order of noun phrases marked for the genitive case cannot be reversed without changing the meaning:

Nie znajdziecie na żadnej z map [pause]***poetów***
NEG find:FUT:2PL on none from map:PL:GEN poet:PL:GEN
tajemniczej ***ojczyzny***
mysterious:SG:GEN fatherland:SG:GEN
'You will not find, on any of the maps, poets of the mysterious fatherland.'

The sequence of the two noun phrases can be moved as a single entity while retaining the same meaning:

(15)
Tajemniczej ***ojczyzny*** ***poetów***
mysterious:SG:GEN fatherland:SG:GEN poet:PL:GEN
Nie znajdziecie na żadnej z map
NEG find:FUT:2PL on none from map:PL:GEN
'The mysterious fatherland of poets, you will not find on any of the maps.'

These constraints on the configuration of the two noun phrases marked by the genitive case are significant in view of the fact that the rich nominal inflectional system of Polish allows the use of various linear orders for a wide variety of pragmatic functions

while retaining the same grammatical and semantic relations among the elements of the utterance. The number of possible orders involving noun phrases with different case marking is n! (n factorial), where n = the number of noun phrases plus the predicate. In the following sentence, n! = 24, and each of the 24 possible orders of the four elements in the sentence yields a grammatical utterance with the same semantic relations among its components:

(16) *Jan* *pija* *wino* *wieczorem*
John:NOM drink:PL:3M:SG wine: ACC evening:INSTR
'Jan drinks wine in the evening.'

When the phrase *tajemnicza ojczyzna poetów* 'the mysterious fatherland of poets' serves as the subject or the object of the affirmative clause, the three components of this noun phrase can be put in any order (3!) and still retain the same meaning. The reason the different word orders are allowed is that the head noun with its adjectival modifier bears the accusative case and the nominal modifier bears the genitive case. Hence, the head and the nominal modifier bear different case marking:

(17)
znajdziecie *na* *każdej* *z* *map*
find:FUT:2PL on each from map:PL:GEN
tajemniczą ***ojczyznę*** ***poetów***
mysterious:SG:ACC fatherland:SG:ACC poet:PL:GEN
'You will find on each of the maps the mysterious fatherland of poets.'

znajdziecie *na* *każdej* *z* *map*
find:FUT:2PL on each from map:PL:GEN
poetów ***tajemniczą*** ***ojczyznę***
poet:PL:GEN mysterious:SG:ACC fatherland:SG:ACC
'You will find on each of the maps the mysterious fatherland of poets.'

znajdziecie *na* *każdej* *z* *map*
find:FUT:2PL on each from map:PL:GEN
tajemniczą ***poetów*** ***ojczyznę***
mysterious:SG:ACC poet:PL:GEN fatherland:SG:ACC
'You will find on each of the maps the mysterious fatherland of poets.'

znajdziecie *na* *każdej* *z* *map*
find:FUT:2PL on each from map:PL:GEN
poetów ***ojczyznę*** ***tajemniczą***
poet:PL:GEN fatherland:SING:ACC mysterious:SG:ACC
'You will find on each of the maps the mysterious fatherland of poets.'

If there is no other noun in the clause, there is considerable freedom of order for the components of the clause. The number of allowed configurations is 4!, with each configuration providing the same referential meaning:

(18)

znajdziecie	*tajemniczą*	*ojczyznę*	*poetów*
find:FUT:2PL	mysterious:SG:ACC	fatherland:SG:ACC	poet:PL:GEN

'You will find the mysterious fatherland of poets.'

tajemniczą	*znajdziecie*	*ojczyznę*	*poetów*
mysterious:SG:ACC	find:FUT:2PL	fatherland:SG:ACC	poet:PL:GEN

'You will find the mysterious fatherland of poets.' (etc.)

The importance of the constraints on the grammaticality of clauses with two nouns marked by the genitive case is that these constraints cannot be explained in terms of the need for case assignment as proposed in the principles and parameters model, since both nouns bear morphological case. They cannot be explained in terms of 'usage' in any sense of the word since the correlations between the forms and meanings are impervious to usage. The only reason that the position must be deployed as a coding means when two noun phrases in the genitive case co-occur in the clause is because of the principle of functional transparency.

3.3.3 Mina: locative predication and the genitive construction

There is a most interesting phenomenon in Mina (Central Chadic) involving genitive constructions and locative predication, whereby the genitive construction has one form in isolation or when serving as the subject or object of a clause and a different form when serving as a complement of a locative predication involving the locative predicator *á*. Here are the facts, whose full description can be found in Frajzyngier et al. 2005.

The modifying construction in isolation or when serving as the subject or object has the form NP *tə́* NP. The relative order of the components indicates which is the head (the first NP) and which is the modifier (the second NP). The modifier can be a noun, a property concept, or a demonstrative:

(19) a.

ngàzù	*wá*	***tú***	*wàl*	*nàŋ*
foot	DEM	GEN	wife	1SG

'That is my wife's foot.' (the high round vowel on *tú* is a product of vowel assimilation to the following labial glide)

ŋ̀kwə̀	***tə́***	*làvéŋ*
goat	GEN	black

'a black goat'

The construction also has the genitive particle *tə́ w*hen used as a subject or object :

(20) *dúwəŋ* ***tə́*** *məd*ìŋ*gwərzé ábə məlá*
back GEN donkey ASSC wound
'The back of the donkey has a wound.'

(21) *kə́ nzlà dúwəŋ* ***tə́*** *məd*ìŋ*gwərzé* (*zà*)
INF cure back GEN donkey (EE)
'He cured the back of the donkey.'

If the two nouns in the modifying relationship function as a locative complement marked by the locative predicator *á*, the modifying relationship is coded by juxtaposition alone and the genitive particle *tə́* is not allowed:

(22) a. *dĭ dĭ á dì tətə ká cìké á*
put put 3SG put 3PL POS all PRED
məɬə́m làkwát
shore river
'He put them all on the shore of the river.'

(22) b. *túl zà á bìŋ kŭ̀l-yî*
leave EE PRED room ancestral spirit-PL
'He went into the house of ancestral spirits and . . . '

(22) c. *á* ***bìŋ wàl*** *nákà wà*
PART room wife REM DEM
'Into the room of that wife . . .'

Compare a clause where the genitive construction is the subject:

(23) ***bìŋ tə́ wàl*** *nákà wà nék*
room GEN wife REM DEM good
'The room of that wife is good/nice.' (elicited)

Two more examples with the genitive construction in the locative complement clause:

(24) a. *nguul kəts lli ki* ***biŋ wal*** *naka waciŋ*
husband gather meat PREP room wife DEM DEM
'The husband took the meat into the room of that wife.'

(24) b. *wal ngən madaraf ɓət skən tu nguul ngən kəts cikee a wtə məməŋ*
woman 3SG favorite take thing GEN husband 3SG gather all PRED house mother:3SG
'His preferred wife took the things of her husband, and put everything in the house of her mother.'

An explanation for the intriguing variation in the genitive construction lies in the

properties of the locative predicator *á*. The predicator *á* is used when the proposition is a locative predication but the verb is not inherently locative. The predicator is local, coding just the type of predication involved. Given the ungrammaticality of clauses with the genitive particle, I propose that the genitive marker *tə́* is also a predicator, albeit of a different variety. Its function is to indicate that the element X (the head) has the properties Y (the modifier). Thus, when *tə́* co-occurs with *á*, the principle of functional transparency is violated, in that it is not clear to what local predication the given form belongs. Once the genitive particle is omitted, the ensuing structure is treated as one noun phrase without an internal predication.

The evidence for the proposed hypothesis consists of several facts. First, in some contexts, the forms *á* and *tə́* are interchangeable. Since *á* is a predicator, the marker *tə́* is therefore also a predicator:

(25) *ƙə̀ tə́ ngìdʼ* ***á*** *təkóŋ*
cow GEN DEM PRED GEN:1PL
'The cow over there is ours.'

(26) *ƙə̀ tə́ ngìdʼ* ***tə́*** *təkóŋ*
cow GEN DEM GEN GEN:1PL
'The cow over there is ours.'

The marker *á* is related to locative predicators having the same form in other Central Chadic languages. The marker *tə́* may be related to the preposition *ta*, which also occurs in other Chadic languages, e.g. Hdi (Frajzyngier with Shay 2002). That preposition may ultimately derive from a verb *ta* 'stop over', attested in some Chadic languages, e.g. in Mupun, (Frajzyngier 1993).

Another piece of evidence that the two forms belong to the same category is that they cannot co-occur in a sequence. The constraint on co-occurrence of the forms is motivated either by the fact that the forms code opposite meanings, and therefore the result is inherently contradictory, or that the two forms code the same meaning, and therefore the result is completely tautological. Given the fact that in some examples the form *á* can be replaced by the form *tə́*, the constraint on their co-occurrence in the sequence results from the fact that they belong to the same category, albeit coding different predications for the same constituents:

(27) **ƙə̀ tə́ ngìdʼ* ***á*** ***tə́*** *təkóŋ*
cow GEN DEM PRED GEN GEN:1PL
'The cow over there is ours.'

A third piece of evidence for the principle of functional transparency is the behavior of the genitive construction in locative complements that are not inherently locative. In such constructions, the locative function of the complement is marked by the preposition *n*. This preposition may co-occur with the genitive marker. The reason why this is possible is because the head of the genitive phrase is not preceded by a predicator.

Consider the following example (28). The first modifying construction, *dəwáŋ kwàykwàyà* 'back of the hyena', has no genitive marker, while the second, *mə́ tə́ kwàykwàyà* 'mouth of the hyena', is marked by *tə́*. The locative predicator *á* cannot co-occur with *tə́*, since the presence of *tə́* would mark the noun *dəwáŋ* as both part of the locative predication and head of the genitive construction. The verb of the second clause, 'put', is inherently locative and so has the same properties as the predicator *á*. However, there is an important difference between the locative complements of the two predicators. The expression *dəwáŋ tə́ kwàykwàyà* 'back of the hyena' is much more inherently locative than the expression *mə́ tə́ kwàykwàyà* 'mouth of the hyena'. The back of an animal is typically the place where the load is carried; the mouth of an animal is not. The evidence for the inherent locative characteristic of 'back of the hyena' is that *dəwáŋ tə́ kwàykwàyà* 'back of the hyena' is not marked by the locative preposition *n*, while the expression *mə́ tə́ kwàykwàyà* 'mouth of the hyena' must be preceded by the locative preposition *n* in a locative predication. Since there is the locative preposition *n*, the genitive head *mə́* cannot be construed as the argument of two different predications and the genitive marker *tə́* is not omitted:

(28)

séy	*ɓət*	*skàŋ-yíi*	*wàcíŋ*	*dzáw*	*dzáw*	*cìké*
so	take	thing-PL	DEM	tie	tie	all
á	***dəwáŋ***	***kwàykwàyà***		*ɓət*	*líjì*	
PRED	back	hyena		take	bridle (F.)	
dəm	*nə*	***mə́***	***tə́***	***kwàykwàyà***		
put	PREP	mouth	GEN	hyena		

'He took those things and attached them all to the back of the hyena. He took the bridle and put it in the hyena's mouth.'

A final piece of evidence that the genitive marker has the properties of a local predicate is provided by headless genitive constructions, where the property concept word is preceded by the genitive marker:

(29)

tì	***lìvèŋ***	*kám*	*à*	*hánə*	*kə*	*ɗál*
GEN	black	TOP	3SG	suit (F.)	INF	do
məsədəf-yîi		*mə*	*nà*	*nákà*	*wàcíŋ*	
devil-PL		REL	of	REM	DEM	
àmmá	***tə́***	***kwèɗék***	*syì*	*ɗám*	*dáy*	
but	GEN	white	COM	good	surpass	

'The black one is convenient as an offering for the devil, as in the past, but the white one is better.'

4 Coding means

Most of the coding means listed in section 3 are described in Frajzyngier and Shay 2003 and are amply documented in the literature. The goal of the present section is to describe one new coding means, non-categorial inflectional morphology, and to propose a new analysis of the coding through linear order.

4. 1 Parsing morphology

Traditional approaches to item and arrangement inflectional morphology share two characteristics: (1) An inflectional morpheme is attached to, or appears on, only one lexical category; (2) for such morphemes there is a one-to-one or one-to-several correspondence between form and function. Thus, the suffix *-z* (orthographic s) added to nouns in English codes plural, and the suffix *d* codes past tense. The genitive case in Polish is added to nouns and it codes the relationships between nouns, the second argument ('object') of negative predications, and the complement of some prepositions. These are instantiations of categorial morphology. A frequent outcome of such morphological coding is a paradigm wherein slightly different forms code the same function across categories such as person or class.

I shall demonstrate here that there also exists a quite different morphological system, whereby the morphological marking can occur on any of a number of lexical or grammatical categories and where the function of this marking is not related to a single meaning. Such morphological marking codes different relationships for different classes of lexical categories. This type of coding is thus non-categorial, in the sense described above. Here is the evidence and an illustration of such a system.

In Wandala, all lexical items, including independent grammatical morphemes, have at least two forms, and a small class of morphemes has three forms. The large majority of lexical items have a form, labeled 'root', that is characterized by the absence of a word-final vowel, and another form consisting of the root + the vowel *a*. Most lexical items and grammatical morphemes exhibit the latter form in clause- or sentence-final position.

The root form occurs on prepositions and spatial specifiers; on auxiliary verbs when they precede the main verbs; on inherently transitive verbs before the object in perfective aspect; on nouns before the adjectives (the default configuration for modifying construction); on all lexical categories before a complement clause; on verbs before adverbs; on interrogative and negative markers when the precede the object.

The root + *a* form is found in subject pronouns when they precede the verb; topicalized noun phrases in clause-initial position; inherently transitive verbs when they precede subjects in the perfective aspect (objects are more natural follow up of transitive verbs); intransitive verbs when they are followed by the object (traditional grammar would consider *a* to be a transitivizing means here); all interrogative and negative markers when they precede the subject; and most nouns in clause-final position. The root + *a* form indicates a type of boundary. Elements to the right of the boundary are not the expected follow-up of the elements to the left of the boundary. Although the form *a* can

be attached to different categories, it is not a clitic (contrary to a suggestion by an anonymous reader) because it does not code a specific semantic function.

Here is the evidence for the functions of the root and root + *a* forms. For the purpose of demonstrating the contrast I have paired, where possible, the same morphemes in different environments.

Prepositions before nouns have the root form:

(30) *yá* *šà-tr-ú* ***gə̀*** *ɮàmá*
1SG speak-3PL-VENT TO population
'I speak to people.'

Auxiliaries that precede main verbs have the root form:

(31) *á* ***də́*** *žàgàdá* *gdzrè*
3SG FUT escape child
'The child will run away.'

Nouns before prepositions:

(32a) *à* *bərdá-n-ú* ***ɠĩn*** *gə* *málm-á-rà*
3SG pull out-3SG-VENT peanuts TO teacher-GEN-3SG
'He pulled out peanuts for his teacher.' (from somebody)

Compare the noun *ɠĩnà* 'peanuts' in clause-final position:

(32b) *à* *də́* *bərdà* ***ɠĩnà***
3SG FUT pull peanuts
'It/he will grab peanuts.' (about birds, boys playing)

Compare also topicalized noun phrases that end in the root + *a* form (the first line of the example represents broad phonetic transcription, the second line represents the underlying forms of morphemes, the third line represents glosses, and the fourth line represents the translation):

(33) *yó ə́lv* *wándàl* ***ŋánnà*** *à* *fyàrà* *málrùwà* [error]
yó *ə́lv* *wándàl* ***ŋánnà*** *à* *f-y-àr-à*
well speech Wandala DEF 3SG put-1SG-ON-PB
mál-rùwà
older brother-1SG
'This Wandala speech is imposed on me by my older brother.'

Verbs before the adverbs are in the root form:

(34) *tá* ***ptsə̀*** *ɮàbè* *á* *də́-m* *gyàkàrà*
3PL return again PRED go-IN cemetery
'They return/are returning again to the cemetery.'

The punctual extension, which has the form *hè* in clause-final position, has only the root form before adverbs:

(35) *tà ptsə-**hə́** ƫàbé á də́-m gyàkàrà*
3PL return-PNCT again PRED go-IN cemetery
'They returned again to the cemetery.'

The negative marker *kà* (clause-final form) has the root form before a preposition:

(36) *mábà kəƫà kùlà wá ájìyù kə̀gìyà*
má bà kəƫà kùlà wá
HYP FOC count calculus COM
*à jì-y-ú **k** gə̀ ìyà*
3SG surpass-1SG-VENT NEG TO 1SG
'If it concerns counting, nobody surpasses me.'

A noun before an adjective has the root form. The position of the adjective after the noun is the default or expected position:

(37) a. *yándà trə́lv gə̀ tàrə́gdzà gyálà nà*
yá ndà-trə́ lv gə̀ tàr ə́gdzà gyálà nà
1SG speak-3PL speech TO 3PL young girl:PL DEM
ántàr gdzàrzálánà də̀ gìyá
*ántàr **gdzàr** zálá nà də̀gìyá*
CONJ child male DEM behold
'I am speaking to the girls and boys, as follows:'

An adjective before noun has the root + *a* form. The position of the adjective before the noun is not expected:

(37) b. *ə́gdzà šóyá làrúusà*
***ə́gdzà** šóy-á làrúusà*
small story-GEN marriage
'a short story of a marriage'

Inherently transitive verbs in the perfective aspect have the root form before an object:

(38) *yò dìkdì zárvàŋánnà kíni*
*yò dyà-k-dy**ì** zárvà ŋánnà kínì*
well know-2SG-know sesame DEF C.FOC
'You know sesame, don't you?'

A verb followed by a pronoun or noun representing the subject of the same clause has the root + *a* form. The function of this alternation is to distinguish between the subject and object of the same clause:

(39) *mákí dà-r-á-**dà** tàrà mùksè žílé*
HYP- go-3PL-GO-go-PB 3PL woman man
'Before they arrived, the woman, the man . . .'

à ***žàgàdà*** *ɬàkàtə́* *tàrè*
3SG ran fellow 3PL
'Their buddy ran away.' (elicited)

Compare the same verb in clause-final position where it ends in the vowel *e*, which characterizes verbs coding movement away from a source:

(40) *mákáfár ndzə̀dàbákà ə́gdzrè də́ žàgàdè*
má ká fá-r ndzə̀dà bá-kà ə́gdzrè də́ ***žàgàdè***
HYP 2SG put-ON force say-2SG child FUT escape
'If you apply force, the child will run away.'

Verbs and other lexical items have the root form before the complement clause. The function of the root form here is to indicate that the pronominal subject that follows the verb is the subject of a different clause:

(41) *tátsə̀tá dúɮàbè*
tá ***tsə̀*** *tá d-úw ɮàbè*
3PL get up 3PL go-VENT again
'They get up, they go there again.'

Object pronouns have the root form before the nominal object:

(42) *yé šà-**k** úyì cùkwá ngùdî*
1SG tell-2SG story small small
'I will tell you a short story.' (*úyà* 'story')

Compare the subject pronoun before the verb, which has the root + *a* form:

(43) *kàdúhùm tù*
kà *dúw hw mtù*
2SG go:VENT outside or
'Did you go anywhere?'

Interrogative pronouns have the root form before the object and the root + *a* form before the subject. The function of this alternation is to distinguish between the subject and object of the clause:

(44) a. *á bàdà-ná* ***wàr*** *kèllù*
3SG flatter-3SG who Kellu
'who flatters Kellu?'

(44) b. *á bàdà* ***wàrà*** *Nábbà/žíl-nà*
3SG flatter who:PB Nabba/man-DEM
'who does Nabba/this man flatter?'

The existence of non-categorial morphology has several implications. One is that it demonstrates the existence of a coding means that is not linked to any specific lexical

category and does not code one specific semantic function. The interpretation of functions coded by the root and root + *a* alternations depends on what kind of morphemes bear the alternation and what kinds of morphemes follow the alternation. The functional distinction between the root and root + *a* form in Wandala is an outcome of grammaticalization of phrase-final form ending in *a*. When deployed in clause-medial position, this form has come to code a variety of functions.

The other implication of non-categorial morphology is that it demonstrates the existence of grammatical coding that lies outside of head-dependent relations and head-dependent marking (cf. Nichols and Bickel 2005), which are at the foundation of the minimalist program.

4.2 A fine-grained analysis of linear order as a coding means

4.2.1 The problems

Linear order, often referred to as 'word order', is most often treated as an independent parameter of language. In Lexical-Functional grammar, word order is treated as one of the means of coding grammatical relations. Although many studies assume the existence of a basic or prevalent word order (Greenberg 1966), we now know that there are languages for which one cannot postulate the existence of a basic word order (Mithun 1992, Dryer 2005). Greenberg, in his seminal study, took the major categories subject, object, and verb as given and studied the orders of these categories without considering the functions of word order. In Dryer's excellent 2005 studies, word order is treated in a similar way, viz. as the configuration of various elements rather than as a coding means. Functions of word order are most often invoked only when the word order that is tacitly considered as basic is changed, mainly to code a variety of pragmatic functions; hence the terms 'left dislocation', 'fronting', and 'right dislocation'. Within the contemporary generative tradition, word order is still taken to be the basic means of representing grammatical relations. Changes in the underlying order are thus taken to be the results of various movements motivated by the need to 'check' various features postulated in higher nodes of representation.

The discussions of word order so far, including those in Frajzyngier and Shay 2003, are not precise enough, and need to be replaced by a more fine-grained analysis of the forms and functions involved. In particular, I propose that the coding means involving linear order consist not of one but rather of at least four distinct means: (a) position; (b) extraposition; (c) relative order; and (d) precedence.

4.2.2 Position

The coding of a grammatical or semantic function by position crucially requires a reference point. The reference point must be identifiable by some phonological, morphological, and possibly semantic characteristics. The position before or after the reference point can then serve as a coding means for a function.

The necessary evidence that a given position in a given language is a coding means is provided when only one function can be coded by this means. This criterion provides a straightforward methodology to assess whether a position is a coding means. Here is an illustration of the criterion.

In English, which is often taken to be a prototypical SVO language in typological studies, the verb is a reference point for the coding of grammatical relations (Frajzyngier and Shay 2003). A more precise analysis, however, is necessary. The position after a transitive verb is indeed a coding means for the category object, as evidenced by the fact that this position cannot be occupied by any other element, e.g. an adverb:

(45) a. Michael Phelps earned his third world championship gold medal Friday night. (*Washington Post*, Aug. 1, 2009)

(45) b. *Michael Phelps earned **Friday night** his third world championship gold medal.

The position before the verb is not, however, a coding means for the category subject, as other elements can occupy this position. A prepositional phrase or a lexical item whose role is inherently marked can be inserted between the subject and the predicate because the position before the predicate is not the coding means for the category subject (examples from British National Corpus):

(46) Instead, the government **on Friday** accepted the ANC argument that the powers of devolved regional governments should be decided by the elected constituent assembly.

265 Mr Baker proposed a meeting with Israel's Foreign Minister, Moshe Arens, and his Egyptian counterpart, Esmat Abdel-Maguid, after Israeli leaders **last Friday** rejected Egypt's invitation to meet a Palestinian delegation in Cairo.

3745 The Confederation of British Industry **yesterday** said MPs' pay must also be pegged.

The insertion of another noun between the subject and the predicate, or before the subject, violates the principle of functional transparency, in that the roles of the two noun phrases preceding the verb cannot be deduced from the formal means employed:

(47) *Water Michael Phelps earned his third world championship gold medal Friday night.

*Michael Phelps water earned his third world championship gold medal Friday night.

In French, the position after the transitive verb is not a coding means for the category object, as evidenced by the fact that it can be occupied by other categories, e.g. by adverbs (all data from CLAPI spoken corpus):

(48) *j'ai intitulé **donc** euh cette communication . . .*
1SG:have entitle therefore eh this presentation
'I therefore gave the title to this presentation . . . ' (extrait de colloque du CERIEP les 5 et 6 février 1998, http://clapi.univ-lyon2.fr/feuilleter.php)

The position before the verb is not a coding means for any specific grammatical relation, as it can be occupied by a variety of categories, including adverbs (= represents

elision of a vowel that would be present in a written form of French):xxxadd examples with adverbs

(49) *j= vous l= dis pa= c= que*
1SG 2PL 3SG say because

j= l'ai testée et absolument ACCENTUE
1SG 3SG:have test and absolutely
vous ne lui arracherez rien
2PL NEG 3SG pull out:2PL nothing
'I am telling you this because I tried her; you are not going to get any information from her.' (http://clapi.univ-lyon2.fr/feuilleter.php)

The fact that an adverb cannot be inserted between the subject and predicate in French does not necessarily mean that the position before the verb is a coding means for the subject. It may well be that the position after the noun, which is marked for gender and number through the system of articles, is the coding position for the verb, which distinguishes the person only in the first and second plural. One of the consequences of the fact that the verb must be preceded by the subject is that topicalization of the subject in French requires the use of the subject pronoun before the verb in addition to any other coding of the subject:

(50) *bon alors euh cette gosse elle n'a= elle*
well then eh this girl 3SG NEG:have 3SG
ne dira jamais rien
NEG say:FUT never nothing
'Well, this girl, she has not, she will never say anything.' http://clapi.univ-lyon2.fr/feuilleter.php

The constraints on adverb placement in English and French have been the subject of numerous studies within generative grammar. Chomsky 1995: 329ff accounts for the difference by postulating different properties of adverbs in the two languages. The analysis proposed here accounts for the differences by postulating differences between more general properties of the two languages rather than differences between the properties of one lexical category.

In Polish, in the order SVO, frequent in speech and writing, or in any other order, the positions before or after the verb are not coding means for grammatical relations, and consequently an adverb may be freely inserted between the noun and the verb and/or between the verb and the ensuing noun:

(51) *Michael Phelps zdobył swój trzeci*
Michael Phelps earn:M.SG:PRF:PAST his third:NOM/ACC
złoty medal w pi_tek wieczór
golden:NOM medal:NOM in Friday evening
'Michael Phelps earned his third world championship gold medal Friday night.'

Michael Phelps w piątek wieczór zdobył

Michael Phelps in Friday evening earn:M.SG:PRF:PAST
swój trzeci złoty medal
his third:NOM golden:NOM medal:NOM
'Michael Phelps earned his third world championship gold medal Friday night.'

Michael Phelps zdobył w piątek wieczór
Michael Phelps earn:M.SG:PRF:PAST in Friday evening
swój trzeci złoty medal
his third:NOM golden:NOM medal:NOM
'Michael Phelps earned his third world championship gold medal Friday night.'

4.2.3 Extraposition

The term 'extraposition' designates a coding means whereby a certain category occupies the position other than the position that codes its function. By definition, coding by extraposition is available only for structures that code some function through position. Since English codes the object function through position after the verb, the extraposition of the object to the position before the subject, most often found in relative clauses, is available as coding means for a pragmatic function:

5_2_0 <786 m> I think
5_2_0 <787 m> firstly
5_2_0 <788 m> that **what character what physical body** you have
5_2_0 <789 m> and to some extent **what character** you possess

In French, the subject can occur to the right of the verb (Grevisse 1991: 1360 ff). It is not clear, however, whether this configuration results from the extraposition of the subject or from the extraposition of the verb. The following is an example from the spoken corpus:

(52) *comme dit sophie quand un des parents est français*
as say Sophie when on of parents is French
'As Sophie says, when one of the parents is French...' (http://clapi.univ-lyon2.fr/feuilleter.php)

In Polish, extraposition is not a coding means because position is not a coding means.

4.2.4 Relative order
Relative order designates a coding means whereby the relative order of two items belonging to the same phrasal or lexical categories codes a specific function: Given X_1 X_2, then X_1 has the function f1 and X_2 has the function f2. We have seen the illustration of such a phenomenon in the case of two nouns marked for genitive case in Polish. I shall provide the evidence for the existence of this coding means in English, Polish, and Japanese.

In English ditransitive constructions, the sequence NP NP codes the sequence Benefactive-Direct object:

(53) CAU 930 This thing gave **me a shock** on my first flight, a back-seat familiarisation flip. (BNC)

1673 As you may recall, last season we failed to get our name drawn from the Vase hat and, remembering past FA Cup fiascos, I gave **the lads the first round Saturday off**.

499 She then gave **Jacob her maidservant**, Bilhah, as a wife, saying that through her, she would build a family.

The importance of the NP_1 NP_2 construction in English is that it differs from the coding of the direct object by the position after the verb. One can resolve this apparent conflict in two ways. One is to interpret the first argument in the benefactive construction, i.e. NP_1, as the direct object, since the term 'direct object' in English is not associated with any specific semantic function. The other is to postulate that the listener faced with the construction NP_1 NP_2 in the position after the verb automatically interprets NP_1 as benefactive. This is the solution I prefer.

Recall that in Polish, when nouns bear different case marking, word order is deployed to code a variety of pragmatic functions. When case distinctions are neutralized and two nouns in the same predication have the same number, Polish deploys relative order as a coding means for grammatical relations, whereby the sequence NP NP codes Subject Object, regardless of the position of the verb, i.e. when the sequence precedes the verb, follows the verb, or is interrupted by the verb:

(54) *więc choć* ***wielu*** ***jego*** ***rówieśników*** *przyjmowało*
even though many:GEN 3SG:M:GEN peers:GEN accept:PAST:IMPF
w podobnych warunkach ciśnieniowych ***dobrodziejstwo***
in similar:GEN conditions:LOC pressure:ADJ benefaction:ACC/NOM
tyrana
tyrant:GEN
'Even though many of his peers under similar pressure accepted the tyrant's benefaction . . .'

The noun phrases *wielu jego rówieśników* 'many of his peers' and *dobrodziejstwo tyrana* 'tyrant's benefaction' can occur in the position before the verb or after the verb, but only in that order.

In Japanese, noun phrases may be marked for their grammatical roles by means of postpositions. However, this is not always the case in natural discourse, where postpositions often do not occur and the relative order of the unmarked noun phrases indicates their respective roles. In the following examples, the first noun phrase is the subject and the second noun phrase is the object (all examples from natural discourse, courtesy of Saeko Ogihara):

(55) *dooshite watashi namae wasurechatterun daroo*
why I-Ø name-Ø have forgotten I wonder
"Why have I forgotten his name?"
(CallHome: evltest: ja_1970: 255.15 258.68)

acchan ne kaminoke kitta.
Acchan (name)-Ø DP hair-Ø cut-PST
"Acchan cut her hair."
(CallHome: ja:0986: 349.33 351.08)

himurosan konsaato suru-n ya-ro
Mr. Himuro-Ø concert-Ø do-NML CPL-Q
'Mr. Himuro will do the concert, won't he?'
(CallHome evltest ja_1690: 316.03 319.34)

Although *ga* is often glossed as a subject maker, it can also code the second argument of a number of predicates, including the second arguments of verbs 'be able to', 'have', 'need', 'hear', 'see', 'understand', and a number of adjectives (Kuno 1973: 90). When two nouns are marked by the same postposition, their different roles are marked solely by the relative order. I do not gloss *ga* in the example below, although Shibatani 2001 considers both instantiations of *ga* to be subject markers, each representing a different semantic relation:

(56) *dare ga kono uta ga utaeru ka?*
who this song sing-can Q
'Who can sing this song?' (Kuno 1973: 88)

4.3 Conclusions regarding linear order

The purpose of this section was to demonstrate that instead of the too-vague concept of 'word order' we should use the much more fine-grained concepts of 'position', 'extraposition', and 'relative order'. Languages may use some or all of these means. Languages may differ with respect to which coding means they employ and also in how they employ them. Thus, although position is a coding means for grammatical relations in both English and French, different positions are actually involved, and these positions code different grammatical relations. Relative order is an important coding means because of its wide applicability. The different linear orders interact with other coding means, viz. inflectional case-marking, adpositions, and coding on the predicate.

5 Areas of further exploration

In languages with a rich system of nominal inflection, elements marked for case can occur in a large variety of linear orders. Such possibilities have been noted in many Australian languages and in Indo-European languages with rich case inflection involving nouns, adjectives, numerals, pronouns, and determiners (cf. Nordlinger 1998). Consider again the situation in Polish when there are five elements in the clause, four of which are

inflected for case. Such elements can be put into 120 configurations without changing the referential relationships between the predicate and noun phrases:

(57) *najpierw* *Pan* *wynalazł* *to* *wszystko*
first you:NOM discover:3M:SG DEM all:ACC/NOM
swoim *naukowym* *umysłem,*
own:INSTR scientific:INSTR mind:INSTR
'You first discovered all of that with your scientific mind.'

The only sequence of two components that cannot be separated in the above clause is the Determiner-Noun sequence *to wszystko* 'all of that'. All of the remaining components can be put into a large number of orders, namely 7! We have no information about the functional differences between these linear orders, and our methodological tools to analyze those meanings are insufficient. It appears, however, **that a given position may code different functions for different categories**. Thus, putting the adverb *najpierw* 'first' in the position after the subject appears to contradict the listener's statement with respect to the order of events. Use of the adverb in clause-initial position, as in (57), does not contradict any previous statement:

(58) *Pan* ***najpierw*** *wynalazł* *to* *wszystko* *swoim*
you first discover:3M:SG DEM all own:INSTR
naukowym *umysłem,*
scientific:INSTR mind:INSTR
'You have first discovered all of that with your scientific mind.'

A systematic analysis of the multiple word orders of elements with the retention of the same referential meaning may involve statements of the precedence of one element over another, combined with the categoriality of elements, and, crucially, a statement about what kind of utterance follows. For example, putting the adverb *najpierw* 'first' at the end of the clause implies that another utterance follows that will clarify what happened afterwards:

(59) *Pan* *wynalazł* *to* *wszystko* *swoim*
you discover:3M:SG DEM all own:INSTR
naukowym *umysłem,* ***najpierw***
scientific:INST mind:INSTR first
'You have discovered all of that with your scientific mind, first, ...'

Putting the adverb before the possessive pronoun *swoim* 'own:INSTR' changes the scope of the adverb:

(60) *Pan* *wynalazł* *to* *wszystko* ***najpierw*** *swoim*
you discover:3M:SG DEM all first own:INSTR
naukowym *umysłem,*
scientific:INST mind:INSTR
'You have discovered all of that first with your scientific mind . . .' (implying the existence of other discoveries obtained by other means)

The characterization of such languages as having 'free word order' or as being 'non-configurational' takes into consideration only the coding of grammatical relations between the predicate and the noun phrases. As the few examples above indicate, such characterization does not take into consideration the possibility that rich variations in linear order code a rich system of functions.

The formal means available in individual languages interact in the coding of functional domains. The principle behind the interaction is that if a certain formal means is used to code a function involving the categories X and Y, where Y may be of the same category as X, the same formal means cannot be used to code another function involving the same categories (Frajzyngier et al. 2002). Hence, for the coding of a different function involving the same categories, a different formal means must be used. For example, given that the linear order NP NP (with the retention of full stress on both noun phrases) in English codes the relationship Benefactive-Object, this means is not available to code a modifying relationship between two nouns. The modifying relationship must be coded by some other means, such as the preposition 'of', the genitive marker *-s*, nominal compounding, or another means. In Jamsay, a verb-final Dogon (Niger-Congo) language spoken in Mali, the relative order of two noun phrases codes the subject and object respectively. Hence, this coding means is not available for the coding of modifying construction, and the language uses other means for this purpose (Heath 2008). In Frajzyngier et al. 2002 it is demonstrated that if a language uses juxtaposition to code equational clauses, it cannot use this means to code modification of one noun by another.

Such correlations raise the interesting question of whether formal means are exploited for various functions haphazardly or whether there is some ordering whereby some functions have, as it were, the right of the first choice. What kind of methodology could be used to answer such a question is an open question. It is very likely that functions that are coded more often have a priority over functions that are coded less often. Moreover, functions that are coded more often appear to use less marked forms than functions coded that are less often. For example, within the domain of modality, the intended truth is coded by fewer formal means than doubt in truth or negation of truth. The coding of relationships between the predicate and noun phrases (in languages that make such a distinction) has a priority over the coding of the relationships between nouns. The findings from English, Jamsay, and the languages cited in Frajzyngier et al. 2002 support this hypothesis.

6 Conclusions

The present study proposes a very brief outline of a non-aprioristic syntactic theory whose main components are:

(a) The principle of functional transparency operating in all languages
(b) The system of functional domains and predications coded in individual languages
(c) The coding means available in individual languages
(d) The interaction of various coding means in the coding of predications

It follows naturally from this approach that languages are similar when they code the same functional domains and sub-domains and have the same formal means of coding. Languages are different if they differ in any of those components.
The proposed approach allows us to (1) account for differences and similarities among languages; (2) discover formal means and functional categories not previously described; and (3) provide more fine-grained analyses of forms and functions already described.

*Acknowledgments
The work on this study was part of a larger project supported by Pays de la Loire Chaire Régionale de Chercheur Etranger. This study benefited from discussion with Erin Shay and represents a follow up of our joint work on syntax. Erin has read the whole paper and made additional comments on the theoretical and editorial aspects of the paper. Her comments made me revise substantially the first version of this study. I am grateful to Amina Mettouchi for critical comments on selected issues and for directing me to the corpus of spoken French. Students in my Fall 2009 Syntax Seminar, Amanda Howard, Nicholas Williams, Steven Duman, Stephen Vihel, and Samuel Perdue, suffered though the presentation of some elements of this study. Their skeptical questions and comments helped very much in sharpening the hypotheses and argumentation. Saeko Ogihara has culled examples from a corpus of spoken Japanese and commented on their grammaticality. I am most grateful to anonymous referees of this study. Their comments prompted me to revise some statements, choose better examples, and otherwise improve this paper. None of the readers or commentators is responsible for the analyses or conclusions of this study.

References
Barlow, Michael and Suzanne Kemmer. 2000. *Usage-based models of language*. Stanford, CA: CSLI Publications.
Bybee, Joan. 2009. Language universals and usage-based theory. In Christiansen, M.H. Chris Collins, and Shimon Edelman. *Language universals*. Oxford: Oxford University Press, 17-39.
Chomsky, Noam. 1995. *The minimalist program*. Cambridge, MA: MIT Press.
Chomsky, Noam and Howard Lasnik. 1991. Principles and parameters theory. In J. Jacobs, Armin von Stechow, and Theo Vennemann. (eds). *Syntax: An International Handbook of Contemporary Research*. Berlin: Walter de Gruyter. Reprinted in Chomsky 1995.
Culicover, Peter W. 1997. *Principles and parameters. An introduction to syntactic theory*. Oxford: Oxford University Press.
Dryer, Matthew. 2005. Order of Subject, Object and Verb. In Haspelmath et al. (chapter 81).
Evans, Nicholas. 1995. *A Grammar of Kayardild*: With Historical-Comparative Notes on Tangkic. Berlin and New York: Mouton de Gruyter.

Falk, Yehuda N. 2001. *Lexical-functional grammar. An introduction to parallel constraint-based syntax*. Stanford: CSLI.

Frajzyngier, Zygmunt. 1991. The de dicto domain in language. *Approaches to Grammaticalization*. Ed. by Elizabeth C. Traugott and Bernd Heine. Volume I Amsterdam & Philadelphia: Benjamins. 219-251.

Frajzyngier, Zygmunt. 2004. Principle of functional transparency in language structure and language change. In Frajzyngier, Zygmunt, David Rood, and Adam Hodges (ed). *Linguistic diversity and language theories*. Amsterdam: Benjamins,

Frajzyngier, Zygmunt. 2008. *A Grammar of Gidar*. Frankfurt: Peter Lang.

Frajzyngier, Zygmunt. (to appear). *A grammar of Wandala*. Berlin and New York: Mouton de Gruyter.

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, 155-98.

Frajzyngier, Zygmunt, and Mohammed Munkaila. 2004. *Grammatical and semantic relations in Hausa: 'point of view' 'goal' and 'affected object'*. Cologne: Koeppe.

Frajzyngier, Zygmunt and Jan Mycielski. 1998. On some fundamental problems of mathematical linguistics. In Carlos Martin-Vide, ed., *Mathematical and computational analysis of natural language*. Amsterdam/Philadelphia: John Benjamins, 295-310.

Frajzyngier, Zygmunt, and Erin Shay. 2003. *Explaining language structure through systems interaction*. Amsterdam/Philadelphia: Benjamins.

Goldberg, Adele. 2006. *Constructions at work. The nature of generalization in language*. Oxford: Oxford University Press.

Greenberg, Joseph H. 1966. Some universals of grammar, with particular reference to the order of meaningful elements. In Joseph H. Greenberg (ed.), *Universals of Grammar*. Cambridge, Mass: MIT Press, 73-113.

Grevisse, Maurice. 1991. *Le bon usage. Grammaire française*. Twelfth edition prepared by André Goosse. Paris: Duculot.

Haspelmath, Martin, Matthew Dryer, David Gil and Bernard Comrie (eds.), 2005. *The world atlas of linguistic structures*. Oxford: OUP.

Heath, Jeffrey. 2008. *A grammar of Jamsay*. Berlin: Mouton.

Jackendoff, Ray. 2002. *Foundations of Language: Brain, Meaning, Grammar, Evolution*. Oxford: Oxford University Press.

Jakobson, Roman. 1959. On linguistic aspects of translation. In Reuben A. Brower (ed.), *On translation*, Cambridge, Mass.: Harvard University, Press, 232-239.

Kemmer, Suzanne, and Michael Barlow. 2000. Introduction: A usage-based conception of language. In Barlow and Kemmer 2000, vii-xxviii.

Kuno, Susumo. 1973. The structure of the Japanese language. Cambridge, Mass.: MIT Press.

Kurcz, Ida, Andrzej Lewicki, Jadwiga Sambor, Krzysztof Szafran, and Jerzy Woronczak. 1990. *Słownik frekwencyjny polszczyzny współczesnej*. Cracow: Polska Akademia Nauk, Instytut Jezyka Polskiego.

Langacker, Ronald. 2000. A dynamic usage-based model. In Barlow and Kemmer 2000, 1-63.

Lasnik, Howard, Juan Uriagereka, with Cedric Boeckx. 2005. *A course in minimalist syntax. Foundations and prospects*. Oxford: Blackwell.
Mithun, Marianne. 1991. Active/agentive case and its motivations. *Language* 67: 510-546.
Mithun, Marianne. 1992. Is basic word order universal? In: Doris L. Payne, ed., *Pragmatics of word order flexibility*. Amsterdam/Philadelphia: Benjamins.
Nichols, Johanna, and Balthasar Bickel 2005. Locus marking: Whole language typology. In Haspelmath et al., Chapter 25.
Nordlinger, Rachel. 1998. *Constructive Case. Evidence From Australian Languages*. Center for the Study of Language and Information.
Shibatani, Masayohi. 2001. Non-canonical constructions in Japanese. In Aikhenvald, Alexandra, R.M.W. Dixon, and M. Onishi (eds). *Non-canonical marking of subjects and objects*. Benjamins: Amsterdam, 307-354.
Svartvik, Jan and Randolph Quirk (ed.). 1980. A Corpus of English Conversations. Lund: Lund Studies in English 82.: Lund University Press
Tesnière, Lucien. 1965. *Éléments de syntaxe structurale*. Paris: Klincksieck.
Tomasello, Michael. 2003. *Constructing a language. A usage-based theory of language acquisition*. Cambridge, Mass. and London, England: Harvard University Press.
Tuwim, Julian. 1975. *Dzieła zebrane*. Warsaw: Czytelnik.

List of abbreviations

1	First person
2	Second person
3	Third person
ACC	Accusative
ADDR	Addressee
ADJ	Adjective
ANAPH	Anaphor
Ar.	Arabic
ASSC	Associative
ATT	Attributive
C.FOC	Contrastive focus
COM	Comment marker
COMP	Complementizer
CONJ	Conjunction
COP	Copula
CPL	Copula
D	Dependent (aspect)
DAT	Dative
DAT.OR	Dative orientation
DED	Deduced reference
DEF	Definite
DEM	Demonstrative
DU	Dual
DUB	Dubitative
EE	End of event marker
EXCL	Exclusive
F	Feminine
F.	Fula (Fulfulde)

FOC	Focus marker
FUT	Future
GEN	Genitive
GO	Goal orientation
H.	Hausa
HAB	Habitual
HYP	Hypothetical marker
IN	Verbal extension ‘in’
INCL	Inclusive
INF	Infinitive
INSTR	Instrumental
INTERJ	Interjection
L(OC)	Locative
M	Masculine
NEG	Negative
NML	Nominalizer
NOM	Nominative
ON	Verbal extension ‘on’
PL	Plural
PNCT	Punctual
POS	Point-of-view of subject
POSS	Possessive
PRED	Predicator
PREP	Preposition
PROG	Progressive
PST	Past
Q	Question
QUANT	Quantifier
REL	Relative marker
REM	Remote previous mention
SG	Singular
SQ	Content question marker (‘specific question)
STAT	Stative marker
TO	Preposition ‘to’
TOP	Topic marker
UNSP	Unspecified
VENT	Ventive

[i] The use of the term ‘predication’ here is similar to its use in symbolic logic.