Nominalization
Verbalization
constraining a typology of
transcategorial operations
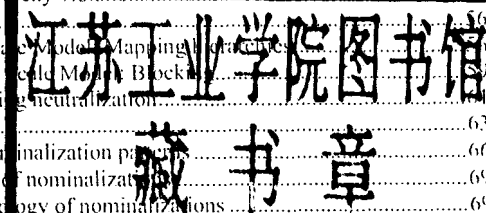
(3)

Nominalization/verbalization

Andrej L. Malchukov

Table of Contents

0. Preface

The present study deals with a typology of transcategorial processes such as nominalization and verbalization. In particular, it addresses the following question: What are the factors that determine the order in which individual verbal categories are lost and nominal categories acquired in transcategorial processes? Although the issue of transcategorial operations in general and nominalizations in particular has been a "hot topic" both in the theoretical linguistics (at least since Chomsky's "Remarks on nominalization") and in the functional-typological literature, these two approaches to this issue have been pursued largely independently. Thus, in a recent analytic overview of the literature on nominalizations, Pieter Muysken states: "So far, no one has succeeded in providing a coherent and unified analysis of action nominalizations, taking into account a sufficiently large sample of languages" (Muysken 1999:252). The present study represents an attempt to provide such an account, as well as to bridge the gap between the formalist and the functionalist approaches to the issue of mixed categories and transcategorial processes.

The study is structured as follows. In §1 I argue for a clear distinction between the two facets of transcategorial processes, decategorization and recategorization. After discussing some earlier proposals aiming to constrain nominalization in §2, I proceed to argue that predictions concerning disruption/acquisition of categories in transcategorial processes are determined by the semantically based hierarchies of nominal and verbal categories, as introduced in §3. In particular, it is argued in §4 that categories that are less relevant to the meaning of the stem are more 'affected' by transcategorial operations. The next two sections present evidence from a sample of languages for this hypothesis demonstrating in turn the gradual disruption of verbal characteristics along the deverbalization cline (in §5) and the gradual acquisition of nominal features along the substantivization cline (in §6). Special attention has been accorded to structural factors (such as morpheme order and category cumulation) that can interfere with the hierarchy constraints. In §7 I propose a 'Generalized Scale Model' for nominalizations mapping the nominal and verbal hierarchy onto each other. Based on the observation that some verbal and nominal categories are conceptually incompatible, I introduce in §8 the notion of 'Blocking' and show how Blocking imposes additional constraints on the Generalized Scale Model. In the next two sections I consider some problematic cases involving Blocking violations and relate the predictions of the

3

Generalized Scale Model to the earlier proposed inductive generalizations concerning nominalization processes. Section 11 addresses the question whether the proposed model is general enough to be applied for other transcategorial processes such as 'verbalizations'. A Generalized Scale Model for verbalizations is proposed, which is a mirror image of the model for nominalizations. As in the case of nominalizations, I first present evidence in support of the denominalization cline and the verbalization cline separately and then address the issues of how these clines relate to each other, as well as consider the Blocking Effects in verbalization processes. In the final section I argue that transcategorial processes are determined by competing – functional and structural   motivations, as well as discuss some further issues related to the role of economy and diachronic factors for the outcome of transcategorial operations. Quite generally, this paper represents an attempt to show how functional and structural factors conspire to constrain the outcome of transcategorial processes.

In general, the approach adopted in this study can be characterized as theory neutral. It follows the tradition of the Leningrad/St.-Petersburg Typology School inasmuch as it attempts to provide a 'calculus' of theoretically available options for a certain grammatical domain, namely for the domain of nominalizations and verbalizations. On the other hand, the present study puts more emphasis on the constraints on the proposed calculus, which is in the spirit of the generative approach. In viewing the constraints as (potentially) conflicting and violable, the approach is similar to the "competing motivations" model in functional linguistics, but also to Optimality Theory. In theoretical assumptions concerning representation of morphological and syntactic categories, the present study is close to the Functional Grammar (as developed by Simon Dik and Kees Hengeveldt) and Role and Reference Grammar (as developed by Robert van Valin), but is also compatible with the recent trends in generative grammar that stress the importance of functional categories for the architecture of clausal and nominal projections. Finally, explanations for the attested generalizations are sought in the functional domain, as is usual in the functional-typological tradition.

1. Introduction: decomposing nominalization

Among transcategorial operations[1] the most studied is nominalization (Comrie & Thompson 1985; Noonan 1985; Lehmann 1988; Dik 1997; Croft 1991; Koptjevskaja-Tamm 1993). Although it is a traditional view that nominalization involves both acquisition of nominal properties and the loss of verbal properties, it has been only recently acknowledged that transcategorial operations such as nominalizations involve both decategorization and recategorization[2]. The term decategorization was introduced by Hopper & Thompson (1984) who showed that verbs or nouns when used not in their primary function tend to lose some of the morpho-syntactic properties associated with their primary function of reporting events and referring to terms, respectively.

For example, when a noun is incorporated it usually loses the ability to distinguish number and to take case-markers and definite articles. Equally important is another facet of transcategorial processes, which was aptly termed by D.N.S. Bhat (1994) recategorization, i.e. when an item used in an extended function acquires some of the properties of those categories to which this function properly belongs.

Thus, a verb used as a referring expression apart from losing some of its verbal trappings, usually also acquires some of the nominal properties such as case, determiners, etc. Hence, traditional terms like "nominalization" actually conflate two properties: "deverbalization" (compare Croft's Deverbalization Scale in Section 2) and substantivization (acquisition of noun-properties)[3].

The most straightforward way to demonstrate the (relative) independence of the processes of decategorization and recategorization is to regard cases where only one is at work. Some arguments for their independence have been already discussed in the literature. As demonstrated by Hopper & Thompson 1984, decategorization need not entail

recategorization: there are examples of incorporation/compounding, which involve decategorization without recategorization. It is important to stress, however, that recategorization need not entail (total) decategorization, either. Thus with regard to nominalizations there are four theoretically possible options (D and R symbols stand for decategorization and recategorization, respectively).

[-D/-R] sentential complements (*I know that he comes*)
[+D/-R] infinitives (*I want him to come early*)
[-D/+R] "clausal nominalizations" (*I disapprove of his driving the car so carelessly*)
[+D/+R] deverbal nouns (*arrival*, etc)

In the first case, involving neither decategorization nor recategorization, we are dealing with a sentential complement. Although it occurs in an NP (Object) position, it retains the structure of a sentence, hardly acquiring any nominal features (e.g., *that*-clauses do not combine with prepositions). In the second case we are dealing with infinitives, which may be viewed as a verb form involving decategorization (deverbalization) without recategorization (substantivization)[4]. In particular, although the infinitive retains VP-internal syntax (e.g., the possibility of taking a direct object) and inflects for aspect and voice, it cannot take tense/mood and agreement markers nor combine with the subject in the sentential (nominative case) form. The third case involving recategorization without (total) decategorization is illustrated here with the gerund in English. Note that the (nominal) gerund combines certain verbal and nominal features; in particular it is capable of taking the ACC object while the subject is expressed as possessor in an NP. Finally in the last case involving both decategorization and recategorization, the verb is completely morphosyntactically assimilated to the (nonderived) nouns, taking all the nominal inflection and retaining hardly any verbal characteristics.

While sentential complements and deverbal nouns representing the opposite poles of the scale are probably universally available options (if one counts cases of paratactic complements as cases of the former and zero conversion as cases of the latter, these options are available even in isolating languages such as Chinese), and the English data are as good as

---

[1] The terms 'transcategorial operation' and 'transcategorial process' are used here in a synchronic sense (the discussion of diachronic issues being relegated to §12.3), as a dynamic metaphor intended to capture the relation between morphosyntactic properties of a derived item and its base (e.g. between a nominalization and a finite verb).

[2] Cf., however, (Walter 1981) for a similar approach. Walter working within Seiler's UNITYP framework characterizes nominalization as involving two funnel operations of increase of nouniness and decrease of verbiness.

[3] Note also the two types of "asymmetry" of finite and non-finite (adverbial) clauses, distinguished by Bisang (2000). The asymmetry can be achieved either by leaving out categories (tense, declarative marker, person) in the subordinate form which are compulsory in the main clause (minus-asymmetry), or by adding markers to the subordinate form (subordination marker, case, switch-reference) which is not necessary in the main clause (plus-asymmetry).

[4] Another case of deverbalization not accompanied by recategorization is found in clause-chaining structures of the "co-subordinate" type, which involve sharing of tense/mood operators (Foley & van Valin 1984). Such cases are not considered here since they do not pertain to the domain of complementation/nominalization.

data from any other language, one can find better examples illustrating the two other intermediate cases. Note that the English infinitive does not represent a case of total deverbalization since it is able to inflect for aspect and voice. In that respect a better example is provided by the "verbal particles" in the Australian language Mangarayi (the term is misleading insofar as the "verbal particles" are not devoid of lexical content). While "verbal particles" have a syntactic distribution comparable to infinitives in that they combine with auxiliaries, they totally lack verbal properties. All verbal categories are either expressed on the auxiliary (tense/aspect/mood and agreement) or (in the case of valence) by auxiliary choice, while the "verbal particles" are uninflected. See the content verbs 'hide' and 'die' in the following example.

(1)   Mangarayi   (Merlan 1982: 15)
        Wurg ø-ga-ni   *ø-waŋgij   jaŋ? wa-o-ma-ñ*
        hide   3SG->3SG-AUX-PAST N.ABS-child die   SUB-3SG-AUX-PAST
        'He hid the child who died.'

On the other hand, verbal particles do not reveal any nominal characteristics and   unlike some other Australian languages – are lexically distinct from nouns (Merlan 1982: 125).

    A better example of the (third) case involving full recategorization without any apparent decategorization is provided by the clausal nominalizations in Hurrian where a clause may be nominalized by the *-še* nominalizer:

(2)   Hurrian   (Speiser 1941: 69)
        ar-oź-au-še-ne-we
        give-PERF-1SG-NZR-SG-GEN
        'of my giving/of my sacrifice'

Note that the verb in (2) retains its finite verbal morphology including tense and AGR suffixes, as well as sentential syntax: the agent cross-referenced by AGR is retained in its sentential (ergative) form (Speiser 1941: 117). On the other hand, the nominalized verb form takes all the appropriate noun inflection including number and case. The case of clausal ([-D/+R]) nominalizations is most interesting for the purposes of my study since it addresses the

question of what combinations of nominal and verbal properties are permitted for nominalizations.

2. Previous research on "category-mixing" in nominalization processes

    As is clear from the previous discussion, processes of decategorization and recategorization are both independent of each other and gradual. For example, a nominalized verb may lose some of its properties (e.g. tense) while retaining some other property (e.g., voice). In a similar way, it may acquire some nominal property (e.g., an ability to take case-markers) but not the other ones (e.g., an ability to distinguish number). Given a large number of verbal and nominal properties the outcome of nominalization processes may be quite diverse. For example, Givón (1990) lists the following operations involved in nominalization:
-       the verb acquires the nominal form
-       verbal TAM marking is either absent or severely restricted
-       verbal agreement is either curtailed or severely restricted
-       case-marking of the subject and object is modified most commonly to the genitive
-       various determiners may be added modifying the former verb   now the head noun; adverbs are converted into adjectival modifiers within the nominalized noun phrase.

    Similar lists of formal adjustments involved in nominalizations can be found in (Croft 1991), (Lehmann 1988) or (Dik 1997). Now the main question is whether there is any ordering of features acquired and lost in nominalization (note that in Givón's presentation these operations are not ordered). Although this question has not been answered in full, there are many proposals to that point in the typological literature (see in particular, Comrie & Thompson 1985; Noonan 1985; Lehmann 1988; Mackenzie 1987; Croft 1991; Koptjevskaja-Tamm 1993; Dik 1997). Thus, (Comrie & Thompson 1985) represents a pioneer study of the typology of (lexical) nominalizations, which has formulated many generalizations that were refined in the subsequent literature. Some of their generalizations are formulated in frequency terms. In particular, they note that aspect and voice may be retained in lexical nominalizations, tense rarely so and mood and verbal-agreement virtually never. Another generalization first formulated in (Comrie 1976) which has figured prominently in the subsequent literature concerns the syntactic typology of nominalizations. It was noted that the subject is the first candidate to receive the possessive (genitive) encoding among the verbal

arguments. That is both S and O may retain sentential encoding or both may be genitivized but if only one argument is genitivized, it will be S while O may retain its sentential marking. All these options may be illustrated by the same English *-ing* form representing - in terms of Ross – a "nouniness squish":

(3) a. *My horse winning the race was no surprise*

  b. *My horse's winning the race was no surprise*

  c. *My horse's winning of the race was no surprise*

In (3a) both S and O retain sentential encoding while in (3c) both are genitivized; in (3b) S is genitivized while O retains its sentential marking. Comrie's syntactic typology of nominalization has been refined and extensively documented by Koptjevskaja-Tamm (1993) for languages of different types (accusative and ergative) as well as languages employing both dependent-marking (i.e. through case) and head-marking (i.e. through agreement) of syntactic relations. Koptjevskaja-Tamm qualifies the (a) pattern of nominalization as sentential (SENT), the (b) pattern as Possessive-Accusative (POSS-ACC) and the (c) patterns as double-possessive, or nominal (POS-POS). Still another frequent nominalizations pattern, also found in English, is Ergative-Possessive (also called Oblique-Possessive). Within this pattern the object is encoded as the possessor, while the subject surfaces as an oblique object (often as an agent in the passive construction).

(3d) *The winning of the race by my (most expensive thoroughbred) horse was no surprise*

Note that this pattern does not provide a counterexample to Comrie's generalization. It still holds: if the object is converted to possessor, the subject cannot retain a sentential form. Thus, Comrie's generalization received overall support, although some counter-examples were also found (see §8 for further discussion).

Many issues raised in (Comrie 1976) and (Comrie & Thompson 1985) have been elaborated on in subsequent literature. Thus, Noonan (1985), Lehmann (1988) and Croft (1991) specifically addressed the deverbalization aspect of nominalization. For example, Noonan (1985: 57) represents the relative susceptibility of verbal categories to be lost in non-indicative complement clauses in the following implication hierarchy:

Subject AGR and Mood > Tense > Aspect > Voice, Valency, Object AGR

This scale reads as follows: the further to the left a category is on the scale the less likely it is to be coded on a non-indicative complement. Croft (1991: 83) proposes the following "deverbalization scale" (> represents entailment relation):

TAM retained > S-marking retained > O-marking retained.

Lehmann's (1988) "Desententialization Scale" is similar in spirit but presents the decategorization process in more detail (> represents a 'prior to' relation).

(a) Constraints on/loss of illocutionary elements > constraints on/loss of mood/modal elements > constraints on/loss of tense and aspect > dispensability of complements > loss of personal conjugation/conversion of subject into oblique > no polarity > conversion of verbal into nominal government > dispensability of subject/constraints on complements
(b) Combinability with adposition/case affix

As is clear from the Desententialization Scale, apart from the loss of sentential properties on the part of the verb, the increasing nominality endows it with distributional properties of a noun (NP), such as combinability with adpositions or case affixes. No particular relation between deverbalization and substantivization processes is stated though.

On the other hand, studies of nominalizations in Functional Grammar (FG) tradition have paid more attention to the ordering of nominal features acquired (Mackenzie 1987; 1996; cf. Dik 1984, 1997). The hierarchy as proposed by Mackenzie and Dik may be represented in a simplified form as follows (> represents an entailment relation):

acquisition of noun features (number/gender, combinability with adjectives) > encoding arguments as genitives/obliques > case marking of nominalization > conversion to non-finite form

Another insight due to Mackenzie is that nominalization involves valency reduction. In a FG account this also explains the ergative alignment of nominalizations, which is found also in accusative languages (see 8.2. for a further discussion of the ERG-POS pattern).

Most recently morphosyntactic "deranking" of the verb in subordinate constructions has been the topic of cross-linguistic study (Cristofaro 2003). S. Cristofaro explicitly addresses the issue of ordering among operations involved in deverbalization and sets up among others the following deranking hierarchies (> represents entailment relation):

loss of aspect > loss of tense
loss of agreement > loss of tense, aspect or mood or expression of TAM categories by special forms
case-marking > loss of agreement and tense, aspect or mood categories

Generally speaking, although the studies within the functional-typological tradition have yielded a number of important inductive generalizations, no principled account has been suggested so far that would allow us to predict the order in which verbal features are lost and nominal categories are acquired. Furthermore, these generalizations are not always compatible with each other. For example, while Lehmann's (1988) or Croft's (1991) deverbalization hierarchies predict that TAM categories are lost prior to AGR, Noonan's (1985) scale for decategorization in complement clauses will predict that tense/aspect will be lost after subject agreement and mood (but prior to loss of object agreement). Another approach to the study of nominalizations represented by the generative tradition may be characterized as deductive and formal (constituency based), rather than inductive and functional. A typical account of nominalizations within the formal approach is the analysis of the English (nominal) gerund by Pullum (1991)[5]. Pullum proposes to analyse the nominal gerund construction as a noun phrase with the verb phrase head along the following lines:

[*my horse's* [*winning the race*]$_{VP}$]$_{NP}$

This account helps to explain why the "mixture" of verbal and nominal characteristics as displayed by the nominal gerund construction is not random. On this analysis the construction is headed by the verb up to a certain level of derivation, which accounts for its verbal properties (e.g., ability to take the direct object). Above this level, however, it is recategorized into an NP, which accounts for its nominal properties (e.g., its ability to take the adnominal

possessor)[6]. Although the proposed mechanism of such recategorization is theory-internal (the author works within the GPSG tradition) the basic insight that the resultant set of verbal and nominal categories depends on which structural level nominalization applies seems to be correct.

Below an attempt will be made to unify different approaches in the study of nominalizations - inductive-typological and deductive-theoretical. On the one hand, the present study of nominalizations seeks to attain typological adequacy in that it is based on a sample of 50 genetically and geographically diverse languages, which are selected according to the Amsterdam-style sampling method of (Rijkhoff et al. 1993) (See the list of the sample languages enclosed in the Appendix). On the other hand, like formal approaches it attempts to give a principled account of constraints on transcategorial processes such as nominalization. In the following sections I shall suggest that transcategorial operations are constrained by hierarchies of verbal and nominal categories, which is reminiscent of some treatments of nominalizations within the formal approach.

3. Theoretical preliminaries: Category hierarchies and transcategorial operations

3.1. Hierarchy of verbal categories and the layered clause structure

In this section I shall claim that predictions concerning the outcome of nominalization processes may be made on the basis of hierarchies of verbal and nominal categories. Such hierarchies have figured prominently in the functional-typological literature. Thus, on the basis of a 50-language sample, Joan Bybee (1985) sets up the following hierarchy of verbal categories:

[[[[[[[V]VAL]Voice]ASP]Tense]Mood]AGR]

According to Bybee, the ordering of categories in the hierarchy reflects the degree of the semantic "relevance" of a given category to the verb stem, that is the extent to which the meaning of an affix directly affects or modifies the meaning of the stem. Thus for example,

---

[5] For an overview of different analyses of the English gerund construction within the generative tradition: see, e.g. (Pullum 1991), (Zucchi 1993).

[6] Cf. (Jackendoff 1977:223): "The point is clear. Gerunds have the constituent structure of sentences up to the X" level and that of NPs above that".

valence or aspect are argued to be more relevant than tense or agreement, since the former affect the meaning of the verb stem more directly than the latter[7].

Empirical evidence for this hierarchy comes from the following cross-linguistic data:
a) Innermost – more relevant – categories are shown to favour lexical/derivational expression, while outermost favour inflectional/syntactic expression;
b) More relevant categories tend to be bound and/or fuse with the stem;
c) The ordering of morphemes, being iconic, cross-linguistically tends to mirror their rank on the semantically based hierarchy (Cf. Baker's (1985) Mirror Principle).

These hypotheses have been tested by Bybee on a sample of 50 languages and received overall support. Thus, with regard to affix ordering her data show that the following ordering preferences obtain in the sample languages, provided that both morphological categories – pairwise compared – are available and both are prefixal or suffixal (Bybee 1985:34-35):
- AGR is external to mood in 13 languages (5 counterexamples)
- AGR is external to Tense in 8 languages (1 counterexample)
- Mood is external to Tense in 8 languages (1 counterexample)
- Tense is external to aspect in 8 languages (no counterexamples)
- Mood is external to aspect in 10 languages (no counterexamples)
- AGR is external to aspect in 12 languages (1 counterexample)

The latter diagnostics that pertains to affix order lends itself most easily for cross-linguistic testing and has been widely used in both functional and generative literature as evidence that verbal categories constitute a hierarchy (see, e.g., Foley & van Valin 1984; Cinque 1999). The following example from Even (Tungusic) exemplifies the iconic affix ordering:

(4)   Even[8]

        ma-v-gara-ri-tan

        kill-PASS-ITER-PAST-3PL

        'have repeatedly been killed'

[7] According to Bybee (1985: 38 ff.) there is a further diachronic dimension in explaining the ordering preferences. Given that stem-affix combinations develop historically from frequently occurring syntactic combinations, both the frequency with which the two words appear together in syntax and their semantic appropriateness for reanalysis reflects the degree of their mutual relevance.
[8] The Even examples cited in the text come from my fieldnotes.

The same ordering preferences have been attested for clitics and auxiliaries, as noted in (Foley & van Valin 1984; Hengeveld 1999; Cinque 1999)[9]. Thus the following English example from (Hengeveld 1999) can be used to illustrate the iconic order of (evidential) mood, (past) tense and (inchoative) aspect in the auxiliary string:

(5)   *The tree must have begun to grow*

The attested cross-linguistic preference for the iconic order of affixes does not imply, however, that in every particular case the order of morphemes mirrors their rank on the functionally based hierarchies (as seems to follow from a (generalized version of) Baker's (1985) Mirror Principle). Firstly, 'low-level' morphological and phonological factors, such as metathesis, can interfere with the iconic order (see, Rice (2000: 361-364) for discussion of these factors with regard to Athapaskan languages). Another important factor, which is not necessarily in line with the functional hierarchies, is diachronic: morpheme order reflects the order of the source items grammaticalizing into affixes (see, e.g., Siewierska (1999) for a cross-linguistic study of the origin of the order of subject and object agreement markers). Yet, as noted by Bybee (1985), the order of the morphemes does not need to reflect an earlier order of words, or the chronological order in which inflectional morphemes develop, since the order can be subsequently changed in accordance with the category hierarchy (as in the Pengo examples discussed below)[10].

The hierarchical structure of verbal categories has also been assumed in Functional Grammar (FG), as developed by S.Dik and K. Hengeveld, and in Role and Reference Grammar (RRG), as developed by R. van Valin and W. Foley, where it is part of the notion of layered structure of a clause (Foley & van Valin 1984; Dik 1991; Hengeveld 1992; Dik 1997; Van Valin & LaPolla 1997). In FG terms, the major layers are predicate, predication, proposition and utterance. (A similar distinction is made in RRG, which distinguishes between nuclear, core, clause and sentence layers, respectively). In both FG and RRG verbal

[9] An important qualification is in order here (Hengeveld 1999): the iconically motivated ordering preferences hold only with regard to items (morphemes, clitics or words) of the same expression format, i.e. among affixes, clitics and auxiliaries separately (for a different view see Cinque 1999).
[10] Rice (2000) presents examples from Athapaskan that demonstrate how a change of meaning of a morpheme conditions a change of its position. Mithun (2001) proposes a plausible diachronic scenario for the functionally conditioned externalization of affixes in polysynthetic languages.

categories are treated as operators, which have certain layers in their scope[11]. Illocutionary force operators such as 'declarative' are assumed to pertain to the utterance layer, (epistemic/evidential) mood operators pertain to the proposition layer, tense operators pertain to the predication layer, and (qualitative) aspect has the narrowest scope pertaining only to the predicate layer. One important insight of these functional models is the correspondence between operators and "satellites", which provide a lexical specification for a given grammatical category. For example, adverbs of manner and frequency are taken to refine distinctions in some languages expressed by corresponding aspectual operators and are correspondingly assigned to the same layer of representation. A simplified version of the FG layered structure is represented below.

[ADVif[ADVmod[ADVasp [V (x) (y) ...] ASP] Tense/Mood]IF]

The innermost layer hosts the predicate and its arguments, which are successively expanded by the predicate layer operators/satellites (aspectuals and manner adverbs), predication layer operators/satellites (tense operators and corresponding adverbs), propositional layer operators/satellites (e.g., evidential markers) and, finally, interpersonal level operators/ satellites (illocutionary force markers).

There are clearly some discrepancies between the relevance based hierarchy, as proposed by Bybee, on the one hand, and the scope-based models such as FG and RRG. For one thing, FG and RRG do not treat AGR markers as operators, since AGR does not enter into scopal relations with regard to TAM categories, while individual TAM categories can be defined with respect to each other in terms of their semantic scope (e.g., tense operators have scope over aspect). Still one may invoke a more general principle of Relevance to account for the ordering of TAM and AGR affixes. Bybee argues that (subject) agreement is less relevant than TAM marking since the former pertains to a participant rather than to the event designated by the verb. Another point is that the choice of subject and the concomitant subject agreement are related – both in synchronic and diachronic terms – to the notion of topicality (cf. Givón's (1976) well-known definition of subject as grammaticalized topic and

agreement as a grammaticalised topic-anaphoric pronoun[12]). Since topicality pertains to the organization of discourse (not unlike "interpersonal layer" markers in FG framework) subject agreement is assigned to an external layer in the layered clause structure[13]. Formulated in positive terms, agreement is relevant to syntax (marking the subject-predicate relation) and to discourse-pragmatics (tracking the most topical argument), rather than to the semantics of the verb. This second dimension of relevance has been explicitly recognized in Hansjakob Seiler's bi-directional conception of relevance (Iturrioz 2001: 541). On this view, the feature hierarchies (as the verbal hierarchy above) are shaped by the two funnel scales of relevance: while the internal categories in the hierarchy (e.g. valency) are (more) relevant to the verbal semantics, external (e.g. agreement, illocutionary force) are (more) relevant to syntax and/or pragmatics. The same two dimensions of relevance are traditionally recognized as functional factors underlying the distinction between derivation and inflection. Derivational categories are assumed to contribute to lexical meaning of an item, while inflectional categories are relevant to syntax (see e.g. Anderson 1992, Plank 1994, Haspelmath 2001).

In any case, the ordering of AGR vs. TAM markers provides straightforward evidence for assigning AGR to an external layer on the hierarchy. Note that its position cannot be (totally) accounted for in diachronic terms (late acquisition of AGR in conjunction with particular word order patterns). As demonstrated by Bybee, there is evidence that when AGR markers are "trapped" in word-internal position due to diachronic processes, they will be later restructured to resume the outermost position (Bybee 1985:40). Bybee illustrates externalization of AGR morphology with the example from Pengo, a Dravidian language. In this language the perfect was formed by the addition of the auxiliary /na/ to the Past tense form, bearing tense and person ending. The subsequent development, which occurred unevenly across the person paradigm, involved encliticization of the auxiliary, resulted in AGR doubling. In the most innovative paradigm, AGR occurs only in the final position.

[11] The term 'scope', as used in this paper as well as in the cited functional-typological literature, pertains to semantic compositionality: An item Z is said to have semantic scope over X and Y if semantics of Z is added to that of X and Y as a unit (see Rice 2000: 24 ff. for discussion).

[12] Compare also Blake's (2001: 131 ff.) view that the subject function within an accusative system has a pragmatic rather than a semantic basis, subject being associated with given or topical information.
[13] Note also that the FG model in the standard form would predict that verbal arguments belonging to the innermost layer would be always retained in the reduced construction even when predication operators and satellites are lost. This is clearly not the case for subject, which is lost prior to adverbs on nonfinite forms. Thus, the layered model, which is helpful in constraining the range of operators/satellites retained in complementation (Hengeveld 1999), cannot account for the subject/object asymmetries attested in nominalizations (Comrie's generalization), unless one assumes that the subject-assignment rules change the rank of the subject argument in the hierarchy. In my view, the standard FG representation treating arguments alike better fits "subjectless" than "subject-oriented" languages (see section 10 on this point).

Schematically this development is represented as follows (for the 1SG form of the verb 'to see'): *huʈaŋ+na -> huʈaŋ-na-ŋ -> huʈa-na-ŋ* 'I see'.

Apart from the position of AGR in the hierarchy there are other points of controversy, such as the relative ranking of tense and mood categories, or the position of negation. Clearly, relative ranking of tense and mood categories will depend on the semantics of mood markers (see below). However, it seems safe to say that with the exception of "root modality", mood markers with epistemic, evidential as well as illocutionary function have a wider scope than tense. With regard to the latter this is uncontroversial, with regard to the former it was demonstrated by Nuyts (2000). He noted, in particular, that modal verbs with epistemic/evidential function (such as English modal *may*) have an idiosyncratic meaning in the past tense, which may be taken as evidence that they cannot appear in the scope of tense operators. The position of negation in the hierarchy is more problematic, since negation operators may differ in scope and pertain to different semantic layers. Indeed, Dik 1997 distinguishes as many as four different types of negative operators differing in scope (see also Zanuttini 1997 for syntactic arguments in favor of multiple positions of negative markers in the clausal structure). This is reflected in a different placement of negative markers if expressed by verbal affixes, both cross-linguistically and within one language. Thus, in Turkish negation belongs to the outermost operators, while in Ika it is the innermost operator hosted on the content verb, and in Eskimo its position is variable depending on scope. Given such scopal ambiguities, (verbal) negation will not be assigned any particular rank in the hierarchy and will be largely disregarded in the present study.

The major source of discrepancies between various hierarchies proposed in the literature is, however, the level of the decomposition of particular categories. Thus, obviously many of the categories represented in the hierarchy allow for further decomposition. Perhaps, this is most obvious with regard to the domain of mood and modality. Here at least four modal levels need to be distinguished, which form a subhierarchy of their own (cf. Van Valin & LaPolla 1997; Dik 1997; Van der Auwera & Plungian 1998; Cinque 1999; Nuyts 2000):

illocutionary>evidential>epistemic>root modality

In a similar fashion one may – in line with (Johanson 1996; cf. Dik 1989; Smith 1991) – propose a further decomposition of aspectual values to distinguish between viewpoint aspect (such as perfective/imperfective), quantitative aspect (iteratives, etc), and inherent aspect (Aktionsart).

viewpoint Aspect > quantitative Aspect > inherent Aspect

Finally, for AGR categories the following (sub-)hierarchy may be proposed (cf. Bybee 1985; Rice 2000)[14].

(VAL) > AGRo > AGRs

The reason for the claim that subject agreement (AGRs) has a wider scope than object agreement (AGRo) is that the latter pertains more directly to verbal valency and less directly to the perspectivizing/topicalizing function[15]. Of course, the relative position of Subject as having a broader scope (c-commanding) than O is a common assumption in constituency-based models. My study will provide further evidence for such hierarchy, referring to cases when AGRs is lost in deverbalization prior to AGRo (cf. Noonan 1985).

The most radical decomposition in the domain of TAM categories was undertaken by Cinque (1999). In his influential study Cinque (1999: 76) sets up the following hierarchy (here presented not in the final most elaborate version!) of "functional heads" (affixes and clitics expressing particular TAM functions). This hierarchy is established on the basis of the mutual ordering of "functional heads", as well as ordering preferences among the corresponding adverbials.

Speech Act > Evaluative M > Evidential M > Epistemic M > Past T > Future T > Irrealis M > Habitual A > Anterior T > Perfect A > Retrospective A > Durative A > Progressive A > Prospective A / Root Modality > Voice > Celerative A > Completive A > Repetitive A > Iterative A

---

[14] By AGR I am referring in the first place to person agreement (which may cumulatively express gender/number distinctions as well), rather than gender agreement as such. The latter is more characteristic of the domain of modifier agreement than of predicative agreement. This would make agreement of verbs in class with its arguments as found in some North-East Caucasian languages a less prototypical case of sentential agreement.

[15] See Rice 2000 (342-346) for arguments that AGRs has scope over ASP, which in its turn has scope over AGRo. Thus, aspect can be analyzed as having scope over objects since it is relevant to the interpretation of the object's number. According to Rice, these scopal relations account for the ordering of AGRs, AGRo and aspect affixes in Athapaskan languages. (Other TAM categories are not discussed by Rice, since they are not expressed by verbal affixes in Athapaskan languages).

Here the decomposition of particular categories reaches its limits since the hierarchy pertains not to grammatical categories as such, but to particular grammemes, members of grammatical categories. Note that according to Cinque, grammemes of the same category may differ in scope (Past having scope over Future, Epistemic over Irrealis, Durative over Progressive, Retrospective over Prospective and Habitual over Iterative), while for semantic reasons they are rather expected to be similar with regard to scopal properties. Therefore I attribute ordering preferences for grammemes of particular categories to other factors, in particular, diachronic ones[16]. Thus, for example, Nenets (Samoyedic), as discussed in 12.1, lacks a structurally coherent tense paradigm: the past tense marker occurring outside AGR, and future tense marker inside AGR. This may be regarded in accordance with Cinque's hierarchy as evidence for Past having scope over Future. Yet, the reason for the different placement of the tense grammemes is historical: the Past marker developed from an encliticized (past) auxiliary verb, while the future tense marker developed from the reanalysis of the imperfective aspect marker.

Thus, in what follows I assume the following hierarchy of verbal (resp. clausal) categories, which may be viewed as a compromise between the different hierarchies proposed in the functional-typological tradition.

[MODi[Sb[MODep[ADVt[ADVman[DO[V]VAL]ASP]TE]Mood]AGR]IF]

On the one hand, I assume with Bybee an "extended" version of the hierarchy including agreement and valency/voice categories. On the other hand, I assume with FG and RRG grammarians a matching relation between grammatical categories (operators) and their syntactic correspondents (satellites). The innermost layer hosts valency/voice operators as well as direct object and object AGR markers. The next layer hosts aspectual operators and adverbial satellites expressing aspectual values (e.g. manner adverbs). The next two layers introduce tense and (epistemic) mood operators with corresponding adverbial satellites (tense and modal adverbs). The two outermost layers introduce subject agreement matched with the clausal subject and illocutionary force markers.

It should be noted that the relevant level of decomposition of verbal categories may depend on the grammatical structure of particular languages. Thus, many languages do not

---

[16] Note also the position of FG grammarians who assume that within a layer the ordering of operators may be variable.

distinguish consistently between tense and epistemic mood categories, which are mutually exclusive and thus may be on structural grounds subsumed under a single category (Xrakovskij & Volodin 1979). In a similar fashion many languages do not consistently distinguish between epistemic and illocutionary modality. Indeed verbal mood is often used to render both functions, as in the case of indicative/imperative opposition, which is probably the most common in mood paradigms. Needless to say, the distinctions are not pertinent (at least at the morphological level) for many (isolating) languages poor in verbal morphology[17]. On the other hand, analysis of languages with the rich morphological structure might require introduction of finer distinctions in a particular semantic domain (see discussion of Korean in 5.2.).

3.2. Hierarchy of nominal categories and the layered NP structure

Jan Rijkhoff (1992; cf. Van Valin & LaPolla 1997) working within the FG framework argued for the following layered structure of the noun phrase (here represented in a simplified form):

[DEM [NUM[Adj[N]CL]NB]DET]

The innermost layer, called qualitative by Rijkhoff, hosts qualitative operators such as "nominal aspect" markers pertaining to individuation such as singulative/collective markers or noun classifiers, as found in a number of languages. In my sample this case may be represented by Miao, which uses classifiers when combined with numerals and/or demonstratives (the bracketed numbers in the examples below indicate tonal patterns).

(6)   Miao   (Purnell 1972: 21)
        pe(43) pen (55) nten (55) na (55)
        two   CLF   book   this
        'these two books'

        The nominal class (gender) markers may also be assigned to that layer, as long as they are expressed by overt markers on the head (as in Bantu languages) and have semantic import rather than being assigned purely on formal (phonological or morphological) grounds (cf.

---

[17] Cf., however, Bisang's (1998) account of the order of particles/auxiliary verbs in isolating languages in terms of Bybee's Relevance Principle.

examples of languages with the overt class marking in 6.1.). The gender assigned on formal grounds is difficult to distinguish from declension class and cannot be located on the semantic hierarchy of nominal categories. Adjectives are satellites within the quality layer. The quantity layer hosts number markers (operators) and numerals (satellites). Finally, the "locality" layer hosts determiners, which may be fully grammaticalized operators (articles) or not completely grammaticalized demonstratives etc. Rijkhoff assigns possessors to the same layer, since they perform functions similar to definite determiners: like demonstratives they help to identify the referent of an NP, albeit in an indirect way — through referring to another identifiable entity, its possessor. Still there is evidence that the possessor is internal to the determiner in the hierarchy. Admittedly this is less obvious for "determiner genitive"[18] languages such as English where a possessor is incompatible with determiners (*this/the my book) than for "adjectival genitive" languages, such as Italian or Russian, which consistently place possessors internal to determiners (cf. Russian eta tvoja kniga lit. 'this your book')[19]. Finally, the nominal hierarchy can be expanded by the "relational" layer, hosting case operators (they may be rendered by adverbial relators or coverbs in languages lacking the case category). Again as in the case of verb feature hierarchy, internal categories on the noun hierarchy (e.g. noun class) are (more) relevant to the meaning of the stem, while external categories (e.g. case, determiners) are (more) relevant to syntax and discourse[20].

Thus one can set up the following hierarchy of nominal categories:

[REL[DEM[GEN[NUM[Adj[N]CL]NB]POS]DET] Case]

A similar hierarchy of nominal inflectional categories has been recently proposed by Lehmann & Moravcsik (2000: 753), who assume the following iconic order among nominal categories.

noun stem -derivational-gender/noun class-number-possessive-determination-case

---

[18] The terms "determiner genitive" and "adjectival genitive" are due to C. Lyons (see Lyons 1999 et passim).
[19] In a more recent publication (Rijkhoff 2002), which is a thoroughly revised and extended version of his 1992 dissertation, Rijkhoff assigns demonstratives to the Location Layer, but argues that articles should be better viewed as belonging to the higher Discourse Layer (Rijkhoff 2002: 227 ff.). Thus, he acknowledges that (some) determiners have a wider scope than possessors.
[20] For extensive argumentation in support of the layered model of NP see (Rijkhoff 1992; Rijkhoff 2002; cf. Langacker 1991; Van Valin & LaPolla 1997).

Much like Rijkhoff, Lehmann & Moravcsik account for such ordering in terms of scope ("operator-operand layering"). They note that gender/class is narrowest in scope "since it is a lexical-grammatical category of the stem that is presupposed but not changed by rules of syntax" (Lehmann & Moravcsik 2000: 753). All the other categories have as their semantic locus either the nominal (number, possession), or even the entire noun phrase (case and determiner); (Lehmann & Moravcsik, ibid.).

The hierarchy of noun categories, unlike the verbal hierarchy discussed above, is still in need of empirical justification (except for the well known Greenberg's Universal Nr. 39 with regard to the ordering of number and case markers). Below I shall present evidence for this hierarchy with regard to affix ordering, which is the most straightforward criterion to apply. My statistics based on the 50-language sample (see the Appendix) yields the following results (note that the count pertains to cases of affixal morphology occurring on the same site of the verb; the position of nominal class markers is disregarded due to their rare occurrence).

case external to DET in 4 languages (the opposite order is found in 1 language)[21]
case external to POS 6 languages (the opposite order is found in no language)[22]
case external to number in 22 languages (the opposite order is found in no languages)[23]
DET external to POS in 2 languages (the opposite order is found in no languages)[24]
DET external to number in 5 languages (the opposite order is found in no languages)[25]
POS external to NB in 5 languages (the opposite order is found in 2 languages)[26]

Thus, in general the data from the sample languages provide confirmation for the proposed hierarchy of nominal categories, as well as provides a robust confirmation for Greenberg's Universal Nr. 39, predicting that case markers occur further from the stem than number markers[27]. On the other hand, the direct evidence for the ordering of POS and DET is

---

[21] Burushaski, Limbu, Iraqw and Koasati vs. Koasati (in Koasati, the semantic cases precede and the syntactic cases follow the determiner suffix).
[22] Eskimo, Hungarian, Hurrian, Iraqw, Mangarayi, Sumerian.
[23] Alamblak, Basque, Burushaski, Eskimo, Ika, Iraqw, Itelmen, Hungarian, Hurrian, Ket, Koasati, Krongo, Limbu, Mangarayi, Nahuatl, Ngiyambaa, Nama, Nivkh, Nunggubuyu, Quechua, Sumerian, Tamil.
[24] Bella Coola, Iraqw.
[25] Burushaski, Iraqw, Koasati, Lango, Limbu.
[26] Abkhaz (the suffixal indefiniteness marker), Hungarian, Iraqw, Lango, Sumerian vs. Hurrian (attributivizer with number distinctions), Mangarayi (inalienable possession marker on some nouns precedes number suffix).
[27] Of course, like in the case of the verbal hierarchy, we are dealing here with cross-linguistic tendencies rather than absolute universals. For example, many Uralic (e.g. Finnish) and Altaic languages (e.g. Tungusic) display a non-iconic order of case and possessive affixes. This does not invalidate the hierarchies assumed in this study.

admittedly rather meagre since "bound" determiners are relatively infrequent and still less frequently found on the same site as possessive AGR markers. Still there is indirect evidence for that ordering: note the placement of AGRpos markers internal to number marking in Hurrian and Mangarayi.

The iconic nominal affix ordering may be exemplified by the following examples from Iraqw (here M1 stands for the linker expressing 1st subclass masculine gender).

(7)   Iraqw   (Mous 1993: 92)
      dawa-ku-éen-ar
      hand-M1-1SG.POSS-INSTR
      'with my hand'

(8)   Iraqw   (Mous 1993: 229)
      hhafeeto-wók-sing
      mats-2SG.POS-DEM
      'these mats of yours'

As in the case of verbal hierarchy the iconic ordering cannot be totally attributed to diachronic factors. Thus Haspelmath (1993) provides evidence for restructuring in a nominal form, referring to cases of externalization of case-markers on indefinite pronouns that are "trapped" in a word-internal position.

As in the previous case, for some languages a further decomposition of categories may be needed. Thus some languages differentiate between inalienable and alienable possession. For such languages there is often evidence for the former to have a narrower scope (for example, in Nungubuyu inalienable possession markers are internal to number markers). Notably, as observed by (Seiler 1983; cf. Siewierska 1998), in many languages alienable possession markers are similar/identical to AGRs markers, while inalienable possession markers correspond to AGRo markers (see (75) from Koasati).

On the other hand, some languages may not consistently differentiate between certain levels. This is true for "determiner-genitive" languages such as English, as well as Abkhaz and Lango for which incompatibility of possessive and determiner affixes has been noted. Just as in the case of verbal hierarchy, depending on the language data one may need either a more decomposed or a more simplified version of the noun hierarchy.

***

It should be emphasized that the idea of hierarchical structure of verbal (resp. clausal) and noun (resp. NP) categories is not theory-internal and pertinent exclusively to FG (or RRG). It is also compatible with those versions of generative grammar, which treat grammatical categories as functional categories "licensing" lexical specifiers (Cinque 1999). The architecture of clausal categories with AGR having scope over tense is accepted in some versions of generative grammar (cf. Chomsky 1995)[28]. In a similar way, the proposed hierarchy of nominal categories is similar to the hierarchy of functional projections within the determiner phrase as suggested by S. Abney (1987) and much subsequent generative literature[29]. Thus the analysis in the following sections may be characterized as theory-neutral, although explanations for the attested generalizations are sought in the functional domain, as is usual in the functional-typological tradition[30].

4. Hierarchy constraints on transcategorial operations: The hypothesis

Now, how can the presented above hierarchies of verbal and nominal categories help to constrain the outcome of transcategorial operations? In my view, this is possible if one accepts the following hypothesis: External categories (resp. layers) on the hierarchies are more readily affected by transcategorial operations than the inner layers. That is, external

---

since hierarchies are driven by functional factors (the principle of scope-relevance) that can be at odds with other factors, in particular diachronic ones (as in the aforementioned Uralic and Altaic languages where we are either dealing with the late development of possessive morphology as compared to case morphology or with the rise of declension from encliticized postpositions inflecting for possession). Thus, I do not assume - under (generalized version of) Baker's (1985) Mirror Principle - that the hierarchy is reversed in the cases when affix order is non-iconic (see Spencer 1992 for discussion of the problems with such an approach). Rather, I view function based universal hierarchies and language particular ordering principles as two independent (and potentially conflicting) factors that jointly determine the outcome of transcategorial operations (see §12.1. below for further discussion).

[28] See Siewierska (1991) for a comparison of representations of clausal structure within the FG and GB traditions. Since then the convergence between these two models seems to increase, which has to do with the "explosion of INFL" assumed in current generative models as well as with the increased interest for typological issues.

[29] Cf. e.g. contributions to Alexiadou & Wilder (1998) which assume, or argue for, the (subparts of the) following general architecture of nominal functional projections: [DetP[PosP[NumP[KindP]]]].

[30] In fact, since the draft version of this paper has been completed several recent studies of mixed categories within the generative tradition came to my attention: the programmatic articles by Bresnan (1997) in the LFG framework and by Borsley & Kornfilt (2000) in the GB framework, as well as two book length studies by Malouf (2000) in the HPSG framework and Alexiadou (2001) in the Minimalist framework. Borsley & Kornfilt's approach is especially similar to the approach adopted in this paper, since it also attempts to derive constraints on "category mixing" from a universal hierarchy of functional categories.

categories are more readily acquired/lost in the process of category-changing. This genaralization, if valid, has a functional motivation: external operators reflect the syntactic and/or pragmatic function of a given lexical item more directly than internal. For example, since the ultimate task of nominalizations is to create a referential expression, the use of determiners, specifying the mode of reference and case markers indicating its syntactic function should be possible for the derived item. The same holds for decategorization processes: disruption of morpho-syntax begins from the external layers (illocutionary force, etc), indicating their former pragmatic function[31].

Below I shall present the data on nominalization from the sample languages in support of this hypothesis. In many cases explicit statements as to the absence of specific categories may be missing in the descriptive grammars, which makes it difficult to present the data from the sample languages in a systematic way. Further, the hierarchies will be simplified in another respect: since I shall largely confine myself to "factive" nominalizations (to the exclusion of subjunctives) I shall not systematically distinguish between tense and mood in non-finite forms (see, however, 5.2. on the gradual loss of modal distinctions in Korean).

Much like Koptjevskaja-Tamm (1993) I shall focus on nominalizations with actional/propositional meaning (rather than agent nominals, result nominals, etc), since they are most similar to verbs in argument structure. However, unlike Koptjevskaja-Tamm, I do not restrict myself to lexical nominalizations (derived nouns with actional meaning) to the exclusion of clausal nominalizations that retain (to some extent) features of the finite verb. Rather lexical nominalizations and clausal nominalizations are viewed as two points on the noun-verb (resp. NP-clause) continuum, on which particular nominalizations will be located depending on the degree of their decategorization/recategorization. Furthermore, the distinction between lexical and clausal nominalizations is difficult to maintain on empirical grounds as well, given proliferation of borderline cases (note, e.g., the controversial status of the 'masdar' nominalizations, as found in Altaic languages).

The role of hierarchies in decategorization and recategorization processes will be first examined separately. Section 5.1. presents the cline of deverbalization, section 6.1. the cline of substantivization, subsequent sections address the question how these two hierarchies relate to each other.

---

[31] A similar suggestion relating (in)accessibility of particular verbal categories to loss in deverbalization processes to Bybee's Relevance has been recently made in Iturrioz (2001) with regard to Huichol (Uto-Aztekan).

5. Nominalization as deverbalization

5.1. The cline of deverbalization

The cline of deverbalization as predicted by my model is represented below. Here I shall concentrate on morphological categories of verbs, disregarding for the moment eventual discrepancies between operators and satellites (cf. 5.4.)[32]. The line indicates the range of the verbal categories retained.

[[[[[V]VAL]TAM]AGR]IF]

1) _____
2) _____
3) _____
4) _____
5)

1). At the initial stage of deverbalization all verbal categories including IF markers are retained on non-finite forms. This situation has been reported, for example, for non-finite tenses in Abkhaz, which are used in different types of subordinate clauses, for example, in complement clauses of indirect speech:

(9)   Abkhaz   (Hewitt 1979: 32)
        S-cò-z-ma hºa dɔ-s-à+z+c'aa-yt'
        I-go-NFIN.IMP-Q (saying) he-it-about-ask-FIN
        'He asked me if I was going'

Note that IF markers are partially retained here (see the question particle retained in the embedded question), but partially lost (e.g., politeness markers in indirect commands are found only in direct quotes with *hºa* 'saying'). A similar situation is attested in Korean quotative forms, which retain illocutionary distinctions while speech-level distinctions expressed through the "sentence-enders" get neutralized (see 5.2.).

---

[32] The distinction between morphological and lexical categories (resp. operators and satellites) may be blurred in the case of clitics and particles. The latter will be taken into account and treated as operators inasmuch as they are obligatory (within a given construction).

2). [*IF] Nominalizations. The first stage of deverbalization when the IF markers are eliminated while sentential AGR is retained have been reported for clausal nominalizations in a number of languages. Generally speaking, the loss of illocutionary markers in subordinate clauses is a widespread if not a universal phenomenon. For example, for Tamil it has been noted that non-finite forms cannot combine with sentence-final clitics with illocutionary and evidential value (Lehmann 1973). In Chinese the interrogative particle is reported to be absent in subordinate clauses (Li & Thompson 1981). Most telling, however, in this respect are languages like Quechua and Nama where IF markers are obligatory (at least in some contexts). Thus, in Quechua the use of "validators" (IF clitics), which is possible in verbal sentences and even obligatory in nonverbal sentences lacking a copula (see (117)), is excluded in subordinate clauses. In Nama the declarative marker, which is otherwise obligatory in finite clauses, is absent in nominalizations and relative clauses.

(10) Nama   (Hagman 1973: 235)

Tiita ke //nãati   kè   ≠'aj hãa 'ii
I   DC this.way PAST think PAST.PFV
'I had thought that way.'

(11) Nama   (Hagman 1973: 235)

Tiita //nãati   kè   ≠'aj hãa 'ii-s
I   this.way PAST think PAST.PFV-NZR
'my thinking/that I had thought that way'

Note that apart from the presence of the nominalizer -s, the nominalization in (11) differs from the finite clause in (10) solely in the lack of DECL marker.

3) [*AGR/IF] Nominalizations. The loss of AGR markers in the presence of tense/aspect markers has been reported for nominalizations in Tamil, Krongo, Fijian, Sediq and Lango. Apart from Tamil these languages take the underlying subject in the possessive form (on Tamil see 5.4.). Consider examples below from Krongo and Fijian. In Krongo, nominalizations are marked by the prefix t-, which replaces the subject AGR prefix of independent conjugation. While a pronominal subject is expressed by encliticized pronouns following the pattern of the "dependent conjugation", a non-pronominal subject is marked by POS case. The verbal categories of tense/aspect and valence/voice are retained though:

(12) Krongo   (Reh 1985: 333)

N-á-tàasà   à?àŋ t-óshó-ó-ko-n-tìi   naama à?àŋ
1,2-PRET-want I   NZR-PRET-cook-BEN-TR-2SG thing   me
'I want (that) you cook for me.' (lit. your cooking for me)'

(13) Krongo   (Reh 1985: 333)

(N-á-tàasà   à?àŋ) t-ìimìnò   kà-Sárrà à?àŋ
(I-impfv.want I)   NZR-impfv.help POS-Sara me
'I want that Sara helped me.' (lit. Sara's helping me)

In Fijian nominal clauses the subject marker is replaced by the corresponding possessive pronoun. Otherwise the clause structure remains unchanged, with all verbal categories, including tense (e.g., past aa/future na in (14)) markers, retained (Dixon 1988: 132):

(14) Fijian (Dixon 1988: 131)

au tadra-a   [a o-mu   aa/na   la'o mai]
1SG dream-TR ART CLASS-your PAST/FUT come here
'I dreamed that you had/will come.'

(15) Fijian (Dixon 1988: 131)

[a   o-dra   dau-.'ada.-vai.-rova]
ART CLASS-3PL habitual-run-PASS-rova
'their habitually running with rova'

Lango lacks a tense category altogether, but nominalizations retain aspect (perfective/habitual/progressive), as well as voice (middle, benefactive, ventive) and object AGR (Noonan 1992: 213). Compare different aspectual forms (expressed by stem modification and/or tonal change) of Lango infinitives/nominalizations (Noonan 1992: 92): à-gikó 'stop smth.' (Perfective), à-gikó 'stop smth.' (Habitual), à-gikkò 'stop smth.' (Progressive).

Arguably the same pattern [TAM/*AGR] is represented by nominalizations in Quechua where (relative) tense distinction are retained but expressed in a modified

participial – form. In the following example the -shka form represents the past participle/nominalizer (note that although the subject agreement is lost, the sentential subject position is retained).

(16) Quechua   (Cole 1982: 33)
      Ñuka-ka Juan kay-pi ka-shka-ta   ya-ni
      I-TOP   Juan this-in be-NZR-ACC think-1SG
      'I think that Juan was here.'

Of course, it is typical for participial forms to retain (some of) TAM distinctions, while losing AGR markers. However, only those cases are relevant for the present discussion where participles may also be used as nominalizations, as in the case of Quechua, or Altaic languages; compare the use of the perfect participle as nominalization in Even (Tungusic) in (61) in Section 6.1. below.

4) [*TAM/AGR/IF] Nominalizations. The next group comprises languages such as Fula, Basque and Abkhaz, which lose both TAM and AGR categories in the course of deverbalization but retain valency (compare also the "infinitive-1" form in Itelmen, as in (164) – (165), which is deverbalized up to the valency layer). This group is not sharply differentiated from the one where a nominalized form loses inflectional aspect, while derivational aspect and valency are retained. For example, in Babungo the imperfective/perfective opposition, which is otherwise obligatory, is lost in nominalization, while "circumstantial" aspect suffixes with ingressive, distributive, diminutive, etc meanings are retained. Thus, in Fula "verbo-nominals" comprising participles and infinitive/nominalization in -ki retain voice oppositions just as verbs (participles additionally retain tenses):

(17) Fula   (Arnott 1970: 372)
      suud-u-ki,   suud-aa-k,   suud-ee-ki
      hide-ACT-INF hide-MED-INF hide-PASS-INF
      'to hide, to hide oneself, to be hidden'; also: 'hiding; hiding oneself; being hidden'

"Verbo-nominals" also retain derivational valency suffixes (such as causative), as exemplified for infinitive in (18) and participle in (19):

(18) Fula   (Arnott 1970: 379)
      dar.nu.ki   mootaaji
      stop.caus.INF cars
      'to stop cars'

(19) Fula   (Arnott 1970: 379)
      ma66.it.id.and̃o   (6e   jol.d̃e)
      open.rev.distr.ben.PART (3PL door.PL)
      'who opened all the doors for them'

A similar case is represented by nominalizations in Basque and Abkhaz. In Basque the -te nominalizations (cf. (35) in section 5.3 below) lose all TAM categories hosted on the auxiliary, but retain valency hosted on the content verb. In Abkhaz the 'masdar' (the verbal noun) loses TAM categories including productive aspectual markers (e.g., iterative *la-*), but retains valency categories (reflexive, causative, "version", reciprocal). Example (20) illustrates the retention of the reflexive prefix:

(20) Abkhaz   (Hewitt 1979: 84)
      A-ç-š-rà   o-yə-taxə'-wp'
      the-self-kill-NZR it-he-want-STAT
      'He wants to kill himself.'

The retention of the VAL layer is, however, partial, since the masdar loses AGRo markers as well as the ability to take a DO in the sentential form. Thus, Abkhaz represents a borderline case between this pattern and the next pattern showing loss of valency.

5) [*VAL/TAM/AGR/IF] Nominalizations. Full decategorization involving loss of all categories including valency has been illustrated above for the "verbal particles" in Mangarayi where the content verb in infinitive function does not inflect for any verbal category (see (1)). A similar situation is found in Ket (Paleosiberian), where an infinitive/verbal noun shows a total elimination of verbal categories. This is remarkable since Ket is a polysynthetic language with verbal categories distributed over 14 prefixal and 3 suffixal slots. Notably, not only inflectional slots of the polysynthetic verbal form are lost, but the derivational valency and aktionsart affixes (causative, resultative, etc) are lost as well