Towards a revision of the lexical subcategorization features

Renate Steinitz
FAS

1. The history of the subcategorization features [V] and [N]

The classification of the four major lexical categories Verb, Noun, Adjective and Preposition[1] by means of the features [V] and [N] with two values has done its job for about twenty years.

Advantages:

- The minimal set of distinctive features allows the wanted discrimination without redundancies,
- It allows reference to similar cross-categorial properties of lexical items and generalizations of so called natural classes. (HAIDER (1993:34ff) defines e.g. the parametrization in the direction of government over the feature [N], that is nouns and adjectives behave the same in respect to the direction of government in German and so do verbs and prepositions, [+N] and [-N] define natural classes.)

Disadvantages:

- The features themselves are not motivated independently and their interpretation is somewhat ad hoc and differs among the various accounts.
- With respect to markedness theory P [-V,-N] is a maximally unmarked and A [+V,+N] a maximally marked category. This is clearly not a desirable description of the hierarchy of marked lexical categories in any respect (if we compare the status of the various lexical categories in one language or between languages).

The dissatisfaction with the classic inventory is very much in the air. I refer to proposals in HALE/KEYSER (1994) and WUNDERLICH (1994) in some footnotes. The idea of interpreting lexical categories in the frame work of markedness is due to Wunderlich in his cited papers.

2. Properties of linguistic entities as candidates for subcategorization features

In my work on non-verbal predication and related topics in STEINITZ (1988, 1990, 1991, 1993, 1994) I consider the property that is responsible for assignment of a lexical category to a lexical item.

Neither the morphology of inflection can do the work, if we want to have a definition independently of specific languages with rich morphology, that is if we claim a universal foundation of categorization. Nor can the morphology of derivation; otherwise, simplex words would not be mentioned. I think morphological marking is an epiphenomenon of more basic properties. It cannot be the syntactic functions either, because they don't map one to one into the lexical categories.

---

1 For the moment I have to ignore a fifth subcategory, that is real adverbs such as *gern, sehr, glücklicherweise* etc. A subclass of adverbs such as *hier, dort, so, dann,* etc are pro-adverbials and described as PPs in STEINITZ (1969) already, in terms of current generative theory they are maximal projections of P, or intransitive Ps. To establish a fifth category adverb a third feature is necessary, and the problem is then how to interpret the eight possible combinations. Later on I will try to incorporate the results of ZIMMERMANN (1988a, 1988b, 1994) in my framework. Zimmermann introduces a third feature [Adv] to account for subcategories such as adverbia and attributiva tantum. For my present purposes I keep to two features and interpret all adverbs as intransitive prepositions. The compromise is somewhat fishy. In section 10 it becomes disturbing, because HENGEFELD (1992) do not take PPs but only manner adverbials into account, which would hardly be intransitive prepositions.

In the UNITYP model "semantic parameters" such as "time-stable:time-instable", "related:relational"(cf. e.g HOPPER/THOMPSON 1984, BROSCHART 1991) motivate the distinction of nouns and verbs. Such a parameter is thought of as a bipolar scale, a continuum between the two complementary poles with segments occupied by prototypical nouns and verbs respectively.

I do not deny the affinity of certain categories of meaning to certain lexical categories. Verbs typically denote time instable entities, that is, activities. Nouns typically denote time stable entities, that is, objects and so on. But this level of concrete semantics is not of much help, because once established in the framework of grammar the grid of categories holds for the untypical cases as well. Derivational forms mostly change from a typical into a nontypical lexical category. The categories are now grammatically based independently of their lexical meaning. (But see my attempt to incorporate the model of semantic parameters in the grammatically based one in section 9). It is likely that all of the mentioned properties take part in lexical subcategorization, but I think they are epiphenomenal rather than basic.

3. Semantic type: lexical-syntactic categories

I suppose Formal Semantics, the **theory of structure of meaning**, to be the base for a grid, in which all syntactic and lexical categories in natural languages arrange. There are two distinct categories with further subcategorization, cf. CHIERCHIA (1985):

(a) **Basic, simple categories**

These categories are mapped onto the ontological universe:

1. **Propositions** refer to situations, their semantic type is **<t>**, e.g. they have a truth value.
2. **Individual terms** (definite descriptions and proper names) refer to entities (object entities, situation entities etc.); they are of type **<e>**.

The standard correspondence in syntax:

- (Independent declarative) sentences (Inflection phrases **IP**s,) correspond to the semantic type <t>.
- (Definite) Determiner phrases **DP**s correspond to the semantic type **<e>.**

(b) **Derived, complex categories**

These categories are n-place predicates <e, t>, <<e,t>e> etc, and they are mapped into properties, states, activities etc.

The standard correspondence in syntax:

1. The set of **lexical categories** and their projections. In German and the other Indo-european languages these are V, N, A, P. In the Semantic Form (SF) of a lexical entry there is no discrimination between the lexical categories, they are all n-place predicates, cf. the entries in (ii), section 5, for illustration.
2. The set of **functional categories** corresponds to special semantic predicates: **I(nflection)** corresponds to an operator over situations, which takes a predicate, selects a section of its extension and localizes it in time. (I ignore for my purposes here the category Complementizer or other functional categories). **D(eterminer)** corresponds to an operator over entities, which takes a predicate and by selecting a part of its extension specifies its reference in the world.

The categories in (a) are basic, primitive, saturated and referential; and those in (b) are derived, complex, unsaturated and non-referential.

4. The basic assumption

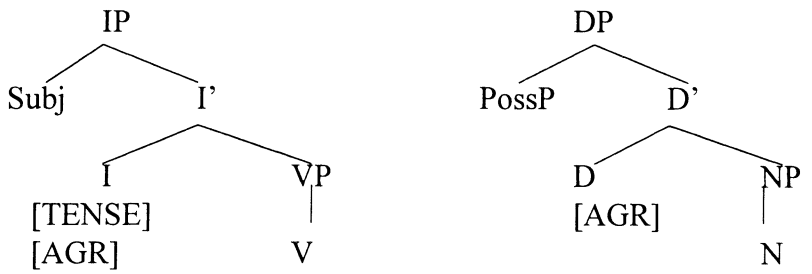
The structures in semantics and syntax are autonomous but systematically related. Every expression in a natural language has to belong to one of the three semantic categories <e>, <t> or a derivation of them. We assume that universally distinct semantic categories are encoded systematically in all languages such that they are distinct at one syntactic level at least. The problem is at which level the distinction has to be fixed.

Referentiality: the basic categories <t> and <e> are both ontological categories, no other category is independently referential. Other categories are referentially dependent predicates. Operators cause the referential anchoring of the predicates, cf. the comments above.

As syntactic correlates to the semantic basic categories <t> and <e> the syntactic phrases IP and DP are the only two categories containing referential arguments. The operation referentially specifying the predicates is **theta-binding**. A functional category I or D as head of IP and DP respectively theta-binds the referential argument of its complement. Theta binding is a selectional constraint, the only complement of I is VP, the only complement of D is NP. The other way round, if there are two distinct lexical categories V and N with a referential argument in a specific language (cf. in (ii) the variable underlined twice in AS) the referential argument of V can only be bound by I and the referential argument of N can only be bound by D.

ABNEY (1986) represents the similarities between sentences (IPs) and Noun phrases (DPs) in the structures (i) (the changes here with respect to other functional categories are not crucial for my point here):

(i)



All projections of lexical categories are semantically predicates, and the functional categories map them onto the basic categories <t> or <e>.

5. Lexical entries in German

The lexicon in the grammar is a model of the lexical knowledge of a native speaker and a lexical entry represents all linguistically relevant information a lexical unit entails. Using the notation of lexical entries e.g. in BIERWISCH (1986, 1988) I give the examples in (ii) for illustration only:

(ii)

PF	LC	SF: AS	SF: PAS
1. ähneln	+V -N	λy λx λw	[w INST [SIMILAR x,y]
2. ähnlich	+V +N	λy λx	[SIMILAR x,y]
3. Ähnlichkeit	V +N	(λy) λx λw	[w INST [SIMILAR x,y]]
4. schlafen	+V -N	λy λx	[x INST [SLEEP y]]
5. schlaf	-V +N	(λy) λx	[x INST [SLEEP y]]
6. schläfer	-V +N	λy	[x INST [SLEEP y]]
7. unter	-V -N	λy λx	[[LOC x] BELOW loc y]

6. The argumentation

(1) In the Semantic Form SF of a lexical entry all lexical units of any subcategory are represented as **category-neutral n-place predicates**. Specific theta roles are assigned to the n arguments, depending on the meaning of the unit.

(2) Argument structure (AS) contains information with respect to the way in which the argument variables are saturated. Internal arguments are saturated by **theta marking** only (if not left unbound). In syntactic terms the arguments get addresses with respect to assigned case, if they don't have structural case, cf. the difference in (iii):

(iii)

	Anna	*(ähneln)*
	(mit) Anna	*(ähnlich sein)*
(die Ähnlichkeit)	*mit Anna*	

(3) The **highest argument** in the hierarchy causes the relevant categorial discrimination of lexical units. It is **a referential or an external argument,** never an internal one.

(a) The hierachically highest argument of the lexical entries 1, 3, 4, 5 and 6 in (ii) is **referential** (underlined twice). Referential arguments are saturated by **theta-binding,** i.e. binding by a functional category I or D. A lexical category with an argument identified as referential in the AS forms a referential category in configuration with a functional category. The **type** of the functional category binding the referential argument determines whether the resulting category refers (via proposition) to a situation or (via individual term) to an object. This is the difference between 1, 4 on the one hand and 3, 5, 6 on the other, i.e. between V and N.

(b) The hierarchally highest argument of the lexical entries in 2 and 7 is **external** (underlined once). Lexical units without referential arguments are referentially dependent, they need a referential anchor to participate in the denotation of a situation or an object. Units of this sort are designated to the **modificational** function. Remaining n-place predications
- they modify situations: adverbal modification; or
- they modify objects: adnominal modification.

A and P have external arguments only and are first of all heads of modifiers. This may surprise the reader, what about the predicative position? I ask him or her for patience until section 8.

Here we have the third kind of saturation of an argument in AS: the external argument of an A or P is **theta-unified** with an argument of the referential phrase, modified by the maximal projection of this A or P.

(4) A referential argument distinguishes V and N from A and P. In STEINITZ (1993), 86 I postulate the feature [Ref]erential to be the first candidate for a revised lexical subcategorization. WUNDERLICH (1993, 1994) incorporates lexical subcategorization into markedness theory. He supposes the complement feature **[Dep]endent** to be unmarked, which then gets the minus value per default in the case of V and N, whereas A and P are [+Dep]-categories. I adopt Wunderlich's proposal.

V and N differ in the **type of reference** they are transmitted to by theta-binding. Candidates for an appropriate feature are the two basic semantic categories <t> and <e> respectively. Again, we have to decide which category is the unmarked one, V or N, and in this case it is not so obvious. In an earlier version of this paper I decided on V to be the maximally unmarked category, probably the only category which appears in every language of the world, I based evidence from SASSE (1988) and others. But the interpretation of the data is not quite convincing yet. For the time being I think that N is the maximally unmarked category. With this I agree with WUNDERLICH (1994) again, whose second feature [Rel]ational as well as the feature [complement] of HALE/KEYSER (1994) gives the same subcategorization with respect to markedness. This is not surprising because we have only the two alternatives. But I do not follow the motivation, I will come back to this in section 11, then I hope to give evidence that only properties I propose are able to motivate lexical subcategorization features. For the moment I take for granted that [<t>] is the second categorization feature, and [-<t>] is equivalent to [+<e>].

Based on the assumption in section 3 and 4 with respect to the basic categories <t> and <e> and their correlates IP and DP in syntax, I conclude that at least these phrases must differ in a grammatically relevant way. That is, V and N have to be differentiated no higher than the projections of the functional heads with which they combine.

The syntactic categories IP and DP are universal categories. The two lexical categories V and N are probably distinct in the default case as well. Note that this assumption is not stated from a European-based view as Sasse supposes, but on the line of my argumentation hitherto.

(5) The modificational and referentially dependent category is not subcategorized so far. In languages with only one modificational category, that is enough. But what to do with clearly distinct categories as the German A and P, although they behave much the same in their syntactic functions? They can both be in adnominal or adverbal position, with a lot of exceptions, which may be relevant, as ZIMMERMANN (1988:284) argues.

I want to take A and P as two multifunctional predicate expressions with certain preferences in their syntactic positions. In the unmarked case A modifies a projection of N, it is theta-unified with the referential argument of N. Therefore A has the features [+Dep], [-<t>]. In the unmarked case P modifies a projection of V, it is theta-unified with an argument of V and has

the features [+Dep], [+<t>].[2] P is determined to be the most marked category, a welcomed result so far.

7. The new feature candidates [Dep] and [<t>] for subcategorization

In contrast to the classic version the base of this subcategorization is not the definiendum itself (V has a V-feature, N has an N-feature etc., whatever these features are). The features are defined independently of the categories they classify, by properties of a kind, which I suppose to be fundamental to the aim of subcategorization.

The features do not rely on specific classes of meaning to which the lexical units belong either, such as 'activity' 'thing' or 'time-stable vs. time-instable' or the like. Nor do they belong to semantic properties such as 'relational'. They rely on that part of AS in SF which collects the formal semantic content, the structure of meaning, and information of its specification in syntax.

In this sense lexical categories are **AS-guided**, that is, they are based in grammar. They are subcategorized with respect to their highest argument in hierarchy. The definition is based on the category specific AS of lexical items, and the specific way of saturation of the highest argument is crucial to the categorization of the items. To summarize:

(iv)	1. [-Dep]:	Theta binding of the highest (=referential) argument:	N, V
	2. [+Dep]:	Theta unification of the highest (=external) argument:	A, P
	3. [-<t>]:	Theta binding by D:	N
		Theta unification with D:	A
	4. [+<t>]:	Theta binding by I:	V
		Theta unification with I:	P

---

[2] Is there empirical evidence for this option? In (i) it is plausible to name the adnominal positon of PP the marked one:

(i) *Helena tanzte durch die Räume*
Die Frau in Troja

It is related to relative clauses (*die Frau, die in Troja war*), in some languages relative clauses are the only option. Other adverbial subcategories are limited to abstract nouns in the adnominal position:

(ii) *die Frau im neunzehnten Jahrhundert / damals*
die Frauen in den Städten / dort
**die Gäste wegen des Geburtstages /* deshalb*
Die Feier wegen des Geburtstages / deswegen
**Das Kind aus Unachtsamkeit /* deswegen*
Die Verletzung des Kindes aus Unachtsamkeit /?deshalb
*Der Besuch aus Vergnügen / * mit Vergnügen*

Situations or events are modified in respect of more different dimensions than objects. With respect to A the adverbial position is the marked one:

(iii) *die schöne Helena*
Helena bewegte sich anmutig

because only "manner adjectives", corresponding to manner adverbials with *-ly*affix in English, can occupy the adverbal position. Other adjectival subcategories are interpreted as predicative attributes:

(iv) *Peter kam krank nach Hause*
Eßt das Obst doch roh

The features reflect the properties lexical units have as lexical heads of syntactic categories in their structural contexts.

It would be near at hand to conclude that lexical categories is mapped onto syntactic positions in a one-to-one fashion, as in (v):

(v)

syntactic position	lexical features [Dep] [<t>]		lexical category
a) subj/obj	-	-	N
b) finite VP (predicate)	-	+	V
c) attributive	+	-	A
d) adverbial	+	+	P

The same features subcategorize the lexical categories and the syntactic positions, it looks like a redundant representation. But the world is not as tidy as that. Lexical category and syntactic position are related, but not identical as supposed in traditional German grammar. The lexical categories can have more than one syntactic function and a function can be met by more than one category, as we have seen in the case of A and P.Thus the situation isn't clear even in German, a language with a specialized system of lexcal categories (cf. HENGEVELD (1992)).

If we take non-Indo-european languages into account the situation seems to be quite confusing, I come back to this topic in sections 9 and 10. And above all, you can have every lexical category except the finite verb in the predicative position.

8. The predicative position

In this context I refer to phrases as in predicative position if they are complements of an overt or covert copula, and I interpret certain configurations of predicative and copula as "**analytic verbs**", which fill gaps in the verbal system. They vary according to the functional perspective of the sentence etc.[3] For illustration see (vi):

(vi)	*schlafen/einschlafen*	:	*wach sein/ wach werden*	(AP + copula)
	zusammenhängen	:	*in Zusammenhang stehen*	(PP + copula)
	---	:	*in Zusammenhang bringen)*	
	--- *ohnmächtig sein*	:	*in Ohnmacht liegen*	
	--- *ohnmächtig werden*	:	*in Ohnmacht fallen*	
	--- ---		*im Zimmer stehen/ sein*	
	--- ---		*ins Zimmer kommen/geraten*	
	regieren	:	*Regent sein*	(DP or NP + copula)
	--	:	*König sein/werden*	
	-- *katholisch sein/werden:*		*Katholik sein/werden*	
	ähneln : *ähnlich sein :*		*Ähnlichkeit haben*	
	--- *ähnlich werden:*		*Ähnlichkeit bekommen*	

A phrase XP in the position **predicative** has following properties:

- it is the complement of a copula,
- it gets no theta-role assignment from this copula,
- therefore, the phrase is not a proper argument, it is not referential,
- but like a normal verb the lexical head of the XP has a theta-marked external argument.
- A succsessfull theta-marking presuppose an argument position already opened by the copula.

A **copula** has the complementary properties:

- it has a referential argument just like a normal verb.
- it has positions for internal and external arguments too, but it does not theta-mark them, thus they turn to pure slots. One slot is filled by the non referential predicative complement and one is filled by the external argument of the head of this predicative complement.

To assign theta-roles to its arguments is a property of syntactic categories corresponding to semantic n-place predicates of type <e,t> , <e<e, t>> etc.; arguments are of type <e>,

---

[3] On the one hand I use the term "predicative" in a restrictive way, only dealing with non referential qualifying predicatives and not with referential identifiing ones, cf. the difference in (i):

(i) *My brother was gardener*
The murder was the gardener.

On the other hand I use the term in a expansive way as the examples in (vi) in the text illustrates.
(a) I subsume the so called periphrastic PPs + copula-like verbs (Funktionsverbgefüge) cf. *in Zusammenhang stehen/ geraten* under the term.
(b) Periphrastic DPs like *Ähnlichkeit haben/ bekommen* have in my opinion the same interpretation, the verb *haben* being a "transitive copula. Cf. the argumentation in STEINITZ (1977)
(c) In STEINITZ (1990) and (1991a) I interpret local and directional adverbials, which are complements of verbs as predicates as well, cf. *im Zimmer sein, stehen/ ins Zimmer kommen, geraten.*
(d) *zu*-infinitivals as *Er hat/ ist zu lieben* probably belong to the predicatives too.

whether they are internal, that is syntactic complements, or external, that is syntactic subjects (in kernal sentences). The copula is the only verb, which has a complement beeing not an argument but a predicate of type <e, t>. The position of the complement can be filled by AP and PP without any complication, they are semantic predicates themselves. What is going on?

In the modificational position A and P have an external argument theta-unified with an argument of the modified phrase, that is both A and P are referentially dependent, they are [+Dep]-categories. In predicative position, however, their external argument has to be theta-marked, just like the external argument of V, a [-Dep]-category. But predicatives need the configuration with a "verb maker" to be real, analytic verbs. The predicative position requires neither a [+Dep]- nor a [-Dep]-head. The feature is irrelevant. In other words, in this position it is not instantiated. That is all that happens in the case of A and P.

But the [-Dep]-category DP can fill the predicative position just as well. In this case, however. the irrelevance of a feature has consequences. In predicative position the DP loses the referential capacity with the syntactic consequence that it gets a lot of constraints with respect to article alternation, modification by relative clauses etc.

The hierachically highest argument of an N in argument position is a referential one, it corresponds to the external argument of an N in predicative position.[4] This is a "process" inverse to the nominalization process in e.g. *lehren* to *Lehrer,* where the external argument of the V corresponds to the referential argument of the N. But in contrast to this process of word formation (and even conversion), in the predicative position we don't have a change of category, an N remains an N in all positions.

The semantic type of the DP in predicative position however has to be <e, t> just like the type of AP, PP, and N projections, which belong to the class of common nouns. I will leave it open for the time being whether we have to describe these facts in terms of "defective DPs"[5] or in terms of "templates", a change in SF only.[6]

What to do with the feature [<t>] of lexical catgories in this position? The normal interpretation according to whcih theta-binding or theta-unification of the highest argument by D or I respectively (see (iv) above) can not be true, because in predicative position the highest argument has to be theta-marked. That is, the feature [<t>], just like [Dep], is not instantiated in predicative position.

To sum up, the predicative position requires no special lexical category. The subcategorization features need not be instantiated. In German all lexical categories can be inserted in this position. The only operation of saturation of an argument is theta marking of its external argument. Corresponding to this the phrases are all category-neutral predicates in SF.

The lexical distinction becomes relevant in non predicative positions:

- In modification position the external argument is unified only [+Dep]-categories can be inserted. Therefore, I called the modificational position the relevant position of AP and PP, distinguishing them from other categories.
- In the position of an individual argument only a [-Dep]-category without an external argument can be inserted, the lexical head has to be an N.

---

[4] You can describe the facts in this way. D loses the capacity of a functional category to bind the referential argument of N. The argument has to be saturated another way; it will be theta-marked, that is the argument type changes and gets a external one.

[5] STEINITZ (1988) is an attempt to interpret predicative nominals as defective NPs.

[6] Cf. DÖLLING (1992) and Blutner (1992).

9.Universality

On this background I can refine a bit what was said in section 2. I insist on the assumption that the grammatical features [Dep] and [<t>] are basic for lexical subcategorization.[7] Are the four lexical categories universal in the sense that they appear in every language as postulated in earlier GB theory? Obviously not.

But there is a minimal condition that every language must meet, this is the result of my work until now:

- Every language has to discriminate between the two semantically basic categories <e> and <t> on some syntactic level; the distinction must be universal.
- All languages are likely to have a distinct category of modification too, adding restrictive or appositive properties to the referent. This distinction too helds not necessarily on the lexical level.

What are the possible combinations and which probably will be realised in natural language? You can see from section 8, that in predicative position the four distinct categories N, V, A and P are redundant even in German, it would suffice to have a non specific category together with a "verb maker".

From that you can deduce the **minimal equipment** a language must have:

- no distinction of lexical categories at the lexical level. This language has n-place predicates as syntactic primes and in addition "functional elements" which specify the predicates.
- a "verb maker" (copula or equivalents) identifies the configuration as the predicate of the sentence.
- a "noun maker" (article, quantifier, classifier ot equivalents) identifies the configuration as the subject or object of the sentence.
- if not combined with a verb- or noun-maker, the external argument of the syntactic prime will be unified with some other argument. The configuration is identified as modification of some kind.

---

[7] My idea on how to reconstruct the obvious affinity of certain categories of meaning to certain lexical categories (cf. section 2) in the framework given is not clear enough. I discuss it briefly in this footnote. Look at the lexical entries 4, 5 and 6 of the sample (ii) in section 5. The three lexical entries have identical representation in PAS, they all are predicates with the form [x INST [SLEEP y]]. They differ in AS and the addresses to be added to their highest, that is referential argument: If bound by I(nflection) the lexical entry is a verb. If bound by D(eterminer) the lexical entry is a noun, the functor D takes a predicate and maps it onto an individual term with the syntactic correlate DP, *das Schlafen, der Schlaf.* In the word formation process the entry loses verbal properties as fixed in the verbal inflection paradigm (tense, mood etc.). Nevertheless the DP refers to the situation entity of sleeping in contrast to the DP *der Mann*, which refers to an object entity. Similarly, a DP in predicative position loses nominal properties, cf. *Peter ist ein Langschläfer*, though the lexical head remains a noun as we have seen.

A lexical entry belongs to a lexical category X according to its feature values. It is a"typical X", if the grammatical properties correlate with the semantic properties in the sense of UNITYP,section 2:

bound by ***semantic category***	***I(nflection): VP***	***D(eterminer) : NP***
situation, event	**typical verb** *schläft*	nominalization *der Schlaf*
property of an object	predicative +cop *ist ein Mann*	**typical noun** *der Mann*

In such a language the features <t> and <e> are not instantiated at a syntactic level higher than X^0.

The combinatory of the two features together with markedness conditions result in a list of possibilities natural language have in respect to their inventory of lexical categories.

(a) **Features are not instantiated**

No instantiation of a feature means no lexical discrimination by this feature, categories collapse.

(vii)

1. Neither [Dep] nor[<t>] is instantiated:	N/V/A/P
2. Only [Dep] is instantiated:	N/V, A/P
3. Only [<t>] is instantiated:	N/A, V/P
4. [<t>] is instantiated in combination with [-Dep] only:	N, V, A/P
5. [<t>] is instantiated in combination with [+Dep] only:	P, A, N/V
6. [Dep] is instantiated in combination with [-<t>] only:	N, A, V/P
7. [Dep] is instantiated in combination with [+<t>] only:	V, P, N/A

The collapse of N/P, V/A is excluded.

(b) **Features are instantiated but not marked** (= underspecified)

The distinction between two categories ist neutralized, the unmarked part of the distinction is generalized. In terms of markedness it means there is no positive value of the feature, the unmarked feature gets the value minus by default. At this point the earlier decision on N as the most unmarked category gets crucial, cf. Section 6.

(viii)

1. Neither [Dep] nor[<t>] is marked, there is one unmarked category only:	N
2. Only [<t>] is marked:	N, V
3. Only [Dep] is marked:	N, A
4. [<t>] is marked only in combination with [-Dep] or	
5. [Dep] is marked only in combination with [-<t>]:	N, V, A,
6. [<t>] is marked only in combination with [+Dep]:	N, A, P
7. [Dep] is marked only in combination with [+<t>]:	N, V, P

3 and 6 are not likely to appear in any language. If there is any lexical discrimination at all, N and V must differ in any case. The constraint to exclude them looks something like:

> If [Det] is marked (has the plus value, the unmarked part gets the minus value by default), then the unmarked part of it must be marked with respect to [<t>].

This is the situation in 5 or 7. You can combine (vii) and (viii), so the possibilities of variation increase. I will leave it as it is.

Is this representation a model to describe the situation in natural languages? Or more precisely, do instantiation and feature-marking form a grid in which differences between languages can be arranged? Although I am more sceptical than in the first version of my paper in respect to the language specific data Hengeveld uses in his book I still think the idea in HENGEVELD (1992) is interesting enough to try a reconstruction in the model I presented here.

10. Hengeveld's typology and its reconstruction in the model

Hengeveld postulates a typology with respect to the lexical categories in various languages (cf. footnote 1):[8]

(A) **Specialized languages** (four lexical categories, normally with four distinct syntactic functions).
(B) **Non-specialized languages** with the subclasses
  (a) **flexible languages** (a single part of speech may be used in different functions)
  (b) **rigid languages** (for certain functions a part of speech is missing)

The sample:

(A) **Specialized** type: German, English etc.: four categories

(Ba) Non-specialized type **flexible**:
  Quechua: **N/A** form one part of speech, which has the syntactic subject/object position of DPs and the position of adnominal modification.
  Lango: **A/Pitr** form one part of speech; the same attributive particle for adnominal and adverbal modification
  Tongan: **V/N/A/Pitr** form one part of speech in all syntactic positions. The units discriminate at a higher level of X projection. If the data are true this would be the minimal equipment I mentioned above!

(Bb) Non-specialized type **rigid**:
  Hausa: **A is missing**, the attributive function is occupied by Ptr ('a person with kindness')
  Wambon: **Pitr (manner adverb) is missing**, medial verbs (verbalized A) have the adverbial function.
  Inuit: **A and Ptr/itr are missing**, V and N only; N combined with modificational suffixes (like German *-lein* 'little') ('child-little, man-old') and nominalized V are adnominal modifiers, N with oblique case affixes are adverbal modifiers.
  Mand.Chinese: **A is missing**, V and N with relativizer (*de*) have attributive function. **Pitr is missing**, reduplicated V has adverbial function.
  Tuscarora, Cayuga: **N, A, Pitr are missing.** "Predication words" or verbs (SASSE (1988)) get different syntactic functions ('(It is) a cat, it is dead' = 'the cat is dead/the dead cat')

I see the problems, the sample is small. Many of the data are not clear neither in HENGEVELD (1992) nor in the papers by typologists. If the pair "rigid/flexible" is relevant at all and if the sample of languages is somewhat representative and if it reflects the data well (many ifs indeed), I can reconstruct "rigid/flexible" in terms of the features [Dep] and [<t>] together with statements about markedness.

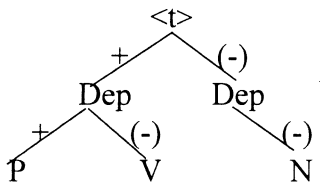
[8] As I already mentioned in footnote 1, I ignored the abdverb as the fifth category for not so good but understandable reason. But now it becomes disturbing, because Hengeveld ignores just prepositions and takes only manner adverbs into account. My comporise in this section is that I name these adverbs Pitr (intransitive P) in contrast to prepositions, which I name Ptr.

Rigid:flexible and markedness

(A) In specialized languages the whole combinatory of two features with two values holds; result: four categories.
(B) In non-specialized languages there are two different ways to take into acount:
  (Ba) In flexible languages at least one feature remain uninstantiated, with the result: less than four categories with mixed properties (two or more of the general possible categories collapsed).
  (Bb) In rigid languages all categorization features are instantiated, but at least one is unmarked and gets the value minus by default. The result is less than four categories, but clear discrimination of each category by two features.

To illustrate this take Chinese, a rigid language in Hengeveld's typology. Chinese is said to have adjectives as a subclass of verbs, cf. GASDE (1993). In the presented model I would say adjectives are not specified at all; Chinese is a language of type (viii-7), where [Dep] is marked only in combination with [+<t>]:

(ix)



In a language with less than 4 categories it is a matter of choice which other category undertakes the job of the "lost" one. There may be preferences, but if Hengeveld is right, the job of the adjective is done:

in Hausa by P,
in Inuit it is N combined with suffixes
in Chinese it is V and N with a relativizer.

11. Is V or N the least marked lexical category?

To sum up, in earlier versions of this paper I thought V to be the unmarked category and I was supported by various descriptions in rigid languages such as Cayuga, cf. SASSE (1988) or Tuscarora in HENGEVELD (1992). They describe these languages as having one category only and they call this category a verb. There is no language with a noun only.

In the meantime I have another idea about the data in Cayuga, even if I have no knowledge of this language, I ask the specialists for discussion. The summary in sections 8 and 9 was that the predicative position can be occupied by a phrase the lexical head of which can be any of the four categories. This position do not need lexical units of any special category, a non specified "predicate" would do the job as well. It might be that the predicate in Cayuga or Tuscarora is such a lexically unspecified entity.

In languages with specified lexical categories any category can occupy the predicative position. In contrast, the argument position, the position of subject or object, can be occupied only by a phrase the lexical head of which is an N. The noun is the only category, which can occupy both the argument position and the predicative position without any change. N is in this sense the least marked category.

Note that my motivation for the decision to classify N as the most unmarked category is quite different from WUNDERLICH (1994). He motivates his decision as follows. N is the most productive category with respect to word formation, almost all other categories can transform into nouns by word formation, it is the most frequent category with the greatest semantic variation, it can be an argument of all other categories, it is learned at the very beginning of the language learning process.

I think most of the arguments are true but too diverse to be relevant criteria. The last argument is relevant but not true, as far as I can see. I suppose the first words a child learns are those like *balla, heia* 'sleep', and *atta* 'walk'. They are category-neutral predicates, i.e. they are used **predicatively**.

HALE/KEYSER (1994) and WUNDERLICH (1994) have similar but not identical subcategorization features. Wunderlich's feature [Rel]ational reflects the assumption that V and P are typically relational, whereas N and A are typically not relational. This feature is based in the PAS of a lexical item, i.e. on meaning categories, more general than the UNITYP-parameters. The feature is also based on semantic features as well. With such features, you must end up with prototypicality. But how does one handle the non-typical cases like intransitive verbs, relational nouns and adjectives? The major motivation of my proposal was to find grammatically based properties for subcategorization, valid for all categories.

HALE/KEYSER (1994) have a more general view on relationality. They interpret the feature [complement] without the notion of typicality. They assume that true verbs are [+complement] or relational units. Verbs behaving differently are taken as results of an incorporation process. So-called unergative verbs like *sneeze* which lack the transitivity alternation are the result of incorporating a nominal complement into an empty verbal head. *Clear* in *The screen cleared* is the result of incorporating an adjectival complement in an empty verbal head etc. The authors denote the structural properties of the four lexical categories with the two features [complement] and [subject], the latter one corresponds to Wunderlich's and my feature [Dep]. Both terms of HALE/KEYSER are syntactic. They suppose a universal system of lexical categories: V is [+complement, -subject], N is [-complement, -subject], A is [-complement, +subject] and P is [+complement, +subject]. In individual languages there are morpholexical processes like incorporation and by this the clarity of the system is often obscured. Though many questions remain open this is a very interesting idea.

12. Invitation

On this basis I want to go further and I need the knowledge of specialists of various exotic languages. I would like to invite them to test the idea to see whether it is supportable. In order to do this it may be necessary to have a new look at his or her data.

Bibliography

Abney,Steven (1986) Functional Elements and Licensing. Glow Conference 86.

Bierwisch, Manfred (1986) On the Nature of Semantic Form in Natural Language. In: F.Klix/H.Hagendorf (eds) Human Memory and Cognitive Capabilities - Mechanisms and Performances, 765-784, Amsterdam-New York: North-Holland Publishers.

(1988) On the grammar of local prepositions. In: Bierwisch, M., W. Motsch, I.Zimmermann (eds)

Bierwisch, M., W.Motsch und I.Zimmermann (eds) (1988) Syntax, Semantik und Lexikon (=studia grammatica XXIX) Akademie-Verlag Berlin.

Blutner, Reinhard (1992) Type shifts in dynamic Montague grammar and the definitness restriction in *there*-insertion contexts. In: Bierwisch, M., W.Motsch und I.Zimmermann (eds), 1 -22.

Bok-Bennema, Reineke (1991) Case and Agreement in Inuit. Studies in generative grammar 38. Foris publications. Berlin, New York

Broschart, Jürgen (1991) Noun, verb, and PARTICIPATION. In: Jakob Seiler/Waldfried Premper (eds.) Partizipation. Das sprachliche Erfassen von Sachverhalten. Gunter Narr Verlag Tübingen, 65-137

Chierchia, Gennario (1985) Formal Semantics and the Grammar of Predication. Linguisitc Inquiry 16, 3 (1995) 417-43

Dölling, Johannes (1992) Flexible Interpretation durch Sortenverschiebung. In: Bierwisch, M., W.Motsch und I.Zimmermann (eds), 23-62.

Gasde, Horst-Dieter (1993) Lexikalische und funktionale Kategorien des chinesischen Satzes. Sprachtypologie und Universalienforschung 46 (1993) 4, 239-274

Hale, Ken, & Jay Keyser (n.d) The limits of argument structure. MIT-paper

Haider, Hubert (1993) Deutsche Syntax generativ. Narr Verlag Tübingen

Hengefeld, Kees, Non-verbal predication. Functional Grammar Series 15, Mouton de Gruyter.

Hopper, Paul J./Thompson Sandra A. (1985): The discourse basis for lexical categories in universal grammar. Language 60/3, 703-752

Riemsdijk, Henk van (1983):The Case of German Adjectives In: F. Heny and B. Richards (eds.) Linguistic Categories: Auxiliaries and Related Puzzles, Vol. One, Reidel Publishing Company, 223-252

Rothstein, Deborah Susan (1985): The syntactic forms of predication

Sasse, Hans-Jürgen (1988): Der irokesische Sprachtyp. In: Zeitschrift für Sprachwissenschaft 7, 2 (1988) 173-213

Steinitz, Renate (1969) Adverbialsyntax. (= studia grammatica X) Berlin: Akademie-Verlag, 206 pp.

(1977) Zur Semantik und Syntax durativer, inchoativer und kausativer Verben. In: Untersuchungen zur deutschen Grammatik I. (= Linguistische Studien des ZISW der Akademie der Wissenschaften der DDR, Reihe A, H. 35) 85-129

(1988) Bewegungenn bei und um Subjazenz, Grenzknoten und Satzglied. In: Bierwisch, M., W.Motsch und I.Zimmermann (eds), 185-218.

(1990) Prädikation, Modifikation und Adverbiale. In: Linguistische Studien des Zentralinstituts für Sprachwissenschaft der AdW der DDR, 206, 117-132.

(1991a) Durative und inchoative Prädikate und die Adverbialkomplemente von Verben. In: Hoffmann, Ludger (Hrsg.), Deutsche Syntax - Ansichten und Aussichten. Berlin, New York: W. de Gruyter. (Jahrbuch 1991 des Instituts für deutsche Sprache, Mannheim). 1992, 186-205.

(1993) Lexikalische Kategorie und syntaktische Funktion. In: Wuppertaler Arbeitspapiere zur Sprachwissenschaft 9, 1993, 79-107.

(1994) Lexikalische Kategorisierung. Talk on the annual meeting of the DGfS
Wunderlich, Dieter (1994a) Lexikalische Kategorien. Talk on the anual meeting of the DGfS.
(1994b) Lexical Categories, Universität Köln, unpublished paper.
Zimmermann, Ilse (1988a) Die substantivische Verwendung von Adjektiven und Partizipien. In: Bierwisch, M., W.Motsch und I.Zimmermann (eds), 279-311.
(1988b) Wohin mit den Affixen? In: Linguistische Studien 179,157-88. AdW der DDR.
(1994) Zur Legitimierung syntaktischer Merkmale. Talk on the annual meeting of the DGfS.