CLIPP

Christiani Lehmanni inedita, publicanda, publicata

titulus
Grammaticography (coautrice Elena Maslova)
huius textus situs retis mundialis
http://www.christianlehmann.eu/publ/grammaticography.pdf
dies manuscripti postremum modificati
04.05.2004
occasio orationis habitae
–
volumen publicationem continens
Booij, Geert & Lehmann, Christian & Mugdan, Joachim & Skopeteas, Stavros (eds.), *Morphologie. Ein Handbuch zur Flexion und Wortbildung*. 2. Halbband. Berlin & New York: W. de Gruyter (Handbücher der Sprach- und Kommunikationswissenschaft, 17/2)
annus publicationis
2004
paginae
1857-1882

170. Grammaticography

1 General problems

1.1 Basic concepts

Like many scientific and, especially, linguistic terms, the term '**grammar**' is ambiguous in designating both a sector of the object area of linguistics and also a scientific account of this sector. Grammar$_2$ is, thus, at a meta-level with respect to grammar$_1$, which is its object. The grammar$_1$ of a language is the systematic part of the way the language maps meanings onto expressions. The grammar$_2$ of the linguist is a scientific representation of this part of a language – in the form of a book, a computer program or some other suitable medium –, whose locus within a comprehensive presentation of a language will be characterized more precisely in 3.

Grammaticography is an activity related to grammar$_2$. This term, too, is ambiguous in the same way. At the object level, grammaticography is (the practice – experience or art – of) grammar writing. At the meta-level, it is the investigation of methodological principles that reconcile this practice with linguistic theorizing. Because of the analogous ambiguity, the term 'lexicography' has been coupled with the term 'meta-lexicography'; and the same could be done for 'grammaticography'. **Metagrammaticography** starts from two ends: On the one hand, it takes stock of successful grammars and distills from them the grammaticographic principles that they follow. On the other hand, it takes successful models of language description as a theoretical basis and deduces from them requirements for an adequate grammar$_2$. Grammaticography is related to investigation of grammar just as lexicography is related to lexicology (investigation of the lexicon).

Metalexicography has had a firm position in (applied) linguistics for several centuries. While grammaticography in the sense of 'production of grammars' goes back to antiquity, metagrammaticography is a fairly recent discipline. The earliest treatment known to us is Gabelentz 1891/1901, Zweites Buch, esp. VI. Capitel: "Die Darstellung der Einzelsprache". The term 'grammaticography' does not appear much earlier than Cherubim 1973. This disproportion between lexicography and grammaticography is not easy to account for. In systematic terms, neglect of grammaticography is simply unjustified and detrimental both to the linguistic discipline and to the quality of actual grammars. In terms of the history of linguistics, there has apparently been a disequilibrium between lexicon and grammar: Research into the former started out as a practical concern, i.e. as lexicography; and only towards the end of the 19th century did lexicology in the modern sense begin. Research into grammar started out as **theory of grammar** with the modists of the 13th century. The beginnings of modern grammaticography with Gabelentz 1891 were taken up by O. Jespersen, but remained largely inconsequential otherwise. Theory of grammar$_1$ got an even stronger position in the second half of the 20th century. The complete neglect of linguistic methodology during the period of the dominance of generative grammar included the complete neglect of grammar$_2$; the concept or problem of grammaticography did not surface in linguistics. As a consequence, most grammars$_2$

actually published, while incorporating conceptions of theories of grammar$_1$ en vogue at their time, do not in their general organization reflect any theoretical foundation at all.

In assessing the relationship between the practice of grammar writing and linguistic theorizing, we may recognize that the last third of the 20th century has brought a significant progress in grammar writing. Apart from many noteworthy monographs, the series *Lingua Descriptive Studies* alias *Croom Helm Descriptive Grammars*, *Cambridge Studies in Linguistics*, *Mouton Grammar Library* and *London Oriental and African Language Library* may be singled out, as they are devoted to comprehensive accounts of hitherto underdescribed languages. This progress in the quality of descriptive grammars can be attributed not so much to some particular model or theory of grammar, but, rather, to a general agreement to keep grammars comprehensible for those outside a particular framework and/or jargon. Roughly since the forties of the 20th century, a certain amount of grammatical descriptions of minority languages had appeared that applied some model of structural or – less often – transformational linguistics. It was soon recognized that these grammars were not even used by linguists, let alone by laymen. Since the seventies of the 20th century, most comprehensive grammars stick to what their authors call "traditional grammar". It was recognized that "traditional" descriptive grammars remain up-to-date and readable much longer than any model-oriented grammars. Such grammars also do not strive for the precision sometimes attributed to formal descriptions. Neglect of precision in itself is not a virtue, but historically it must be seen as a reaction to a kind of hollow precision that served to conceal lack of insight.

In theory, grammaticography and theory of grammar inform each other. In practice, mutual awareness leaves much to be desired. Still, good grammar writing does apply the achievements of linguistic theorizing. Some new concepts and approaches (or more elaborated old ones), presentational techniques, empirical domains to be looked at and included in a grammar, etc. have established themselves as part of recent grammaticographic tradition. To the extent that these innovations are picked up from various models of grammar, this approach is eclectic. Grammars have to be both consistent and comprehensive, and these goals are sometimes in conflict. The **eclecticism** often practiced in contemporary grammaticography may appear as a symptom of inconsistency, but is more properly seen as a consequence of the fact that none of the available theories of grammar suits all languages and all empirical domains equally well. Natural languages comprise heterogeneous phenomena which favor different models and methods. If a grammarian is able to choose the most powerful approach for each descriptive domain and to combine them in the compilation of a comprehensive grammar, this is a virtue rather than a vice. Moreover, this approach to writing grammars is actually the only way to determine the borders of applicability of theoretical concepts and models and their relations to one another, and thus to achieve theoretical progress. In this sense, metagrammaticography provides an interface between the practice of grammar writing and the theory of grammar. On the other hand, metagrammaticography can be viewed as a sort of self-reflection inside the field of descriptive linguistics, which can help to accumulate positive experiences and most suitable (technical) solutions to be applied in future grammatical descriptions.

Like most disciplines concerned with the practical application of scientific insights, and like metalexicography in particular, metagrammaticography is at least as much a prescriptive discipline as it is a descriptive one. It is descriptive insofar as it is faced with

actually published grammars and with the esteem these are held in, a datum which it has to take into account lest it become escapist. It is, however, prescriptive in that it converts both the principles distilled from grammaticographic practice and the theorems of grammatical theory into instructions of grammar writing (cf., e.g., Mosel 2003, Noonan to app.). Metagrammaticography is part of the methodology of linguistics. Consequently, the present treatment will account for observable practice, but will also derive recommendations from it.

1.2 Universalism and particularism

The description of a language has two opposite tasks: to bring out the uniqueness of this language and to render it comparable with other languages. To achieve the former task, the description must not identify the categories and operations of the language with those of other languages, but rather describe the language in its own terms. To achieve the latter, the description has to characterize the language in terms of a linguistic type, presupposing and referring to universal parameters of human language as a background against which the peculiarities of the language stand out. To do this, it must describe the language in the same terms as other languages. The two tasks seem, thus, irreconcilable. Consequently, many existent grammatical descriptions sin by an overemphasis on either **particularism**, which renders the language different from anything that may be familiar to the reader and, thus, unintelligible, or on **universalism**, which converts the language into just another instance of something well-known to the reader and, thus, uninteresting. Both extremes miss their target (cf. also 1.6.1). Many of the grammaticographic principles discussed below revolve around this problem.

The purpose of a language description is not only to tell the reader the facts to any desired degree of detail, but, at a more general level, to convey to him an impression of what this language is like. This evaluation proceeds against the general background of 'le langage', shared by author and reader, and possibly against the comparative background of other languages, too. In the former respect, the grammar will take for granted many aspects of the language that it shares with all other languages (e.g. the sheer existence of word formation, and its most general cognitive and communicative functions). In the latter respect, many grammars, especially textbooks, even presuppose as deserving no comment those features of the language that it shares with the metalanguage of the description. Thus, few grammars of French will bother to mention that that it lacks infixation or that clause structure follows the accusative model of fundamental relations.

1.3 Purpose and kinds of grammar

Grammars may be classified according to a variety of parameters. The first distinction is between a comparative grammar and a grammar limited to one language. There are various kinds of comparative grammars according to the purpose and method of comparison: a **general comparative grammar** is a systematic survey of grammatical phenomena in the languages of the world. A **historical-comparative grammar** traces the evolution of the grammar of a proto-language into its daughter-languages by comparing these. A **contrastive grammar** compares grammatical phenomena of two languages, mostly with some practical application in mind. In contrast to comparative grammars, grammars

devoted to just one language are sometimes called descriptive grammars; but of course such a grammar can also be prescriptive; see below.

A grammar of one language may take the synchronic or the diachronic perspective. A grammar that does the latter is a **historical grammar**. In contradistinction to a language history, a historical grammar is normally not subdivided according to stages of the language, but may rather have the same structure as a synchronic grammar. A historical grammar usually defines one stage of a language – not seldom an ancient stage – as the end point of its scope and differs from a synchronic grammar in tracing the properties of that stage back to the relevant proto-language.

Focusing now on non-comparative synchronic grammars, we may distinguish between what has been known traditionally as **textbook** vs. **reference grammar**. A textbook (German *Sprachlehre*) aims at teaching the language. It arranges grammatical material in the sense of some didactic progression; and grammatical information may be intertwined with other didactic material such as lessons, cultural information and exercises. A reference grammar, on the other hand, gives a systematic and comprehensive overview of the grammatical system or even the whole system of a language, but is limited to that. Its purpose is to provide orderly access to information on the language system rather than to help the user learn the language. It may presuppose some familiarity with the language or with linguistics. There also used to be a distinction between linguists' vs. learners' ("scientific" vs. "pedagogical") grammars. This has lost somewhat in significance because it has been recognized that writing for colleagues is compatible with the requirement of user-friendliness (cf. also the notes on recent developments in 1.1).

The orientation of a grammar may be descriptive or prescriptive. A **descriptive grammar** is based on data that are independent from the analyst and describes these. Its purpose is to ascertain and systematize the facts in some empirical domain. A **prescriptive grammar** represents what the author considers correct. Its purpose is to orient the reader towards the norm of the language.

While the distinction between descriptive and prescriptive linguistics is clear in principle, there are some factors that tend to blur it in practice. Firstly, no grammar (descriptive or other) can account for all the variation that occurs in its object area. It will exclude at least speech errors, false starts, hesitation phenomena etc.; and it may exclude youth language or archaic ritual language from its scope. This delimitation amounts to the identification of a norm. This does not necessarily coincide with the highest norm in the entire speech community; it simply means the identification of a prototypical manifestation of the variety under study and the marginalization of other varieties. The exclusion of some facts from the variety being described is similar to the decisions made in a prescriptive grammar. Secondly, most users of a grammar, even of a descriptive grammar, require reliable information on what is grammatical in the language or, at least, what is normally said. They want to be given rules. To satisfy this need, the author has to distinguish manifestations of the norm from deviations from it and has to formulate rules for grammatical constructions. This, again, is what a prescriptive grammarian does.

In this article we will concentrate on descriptive reference grammars. However, a good portion of the following metagrammaticography applies to other types of grammar as well.

1.4 Data and variation

Data are the empirical basis of a grammar (see Lehmann 2004[D]). There are different kinds of them, and they play different roles in grammars. This is determined, first and foremost, by the orientation of the grammar in the sense of 1.3. In a descriptive grammar, data function as the empirical basis and the object of research. Therefore the data must exist independently from the production of the grammar. A descriptive grammar is therefore based on a corpus of data. In a prescriptive grammar, on the other hand, data have the function of illustrating the norm. They may be taken from a corpus of texts that is conceived as representing the norm – for instance the works of classical writers –; but if the author considers herself master of the norm, then she may simply make up illustrative examples, which are then no data in the strict sense. Some descriptive grammars, it is true, are also based on material produced by the author of the grammar. In this case, however, the data are not independent from the scientist, which means they are not reliable data by scientific standards.

The **corpus** on which a descriptive grammar is based may essentially be of two kinds. It either exists prior to the grammar as a body of texts available in the speech community; or it is collected by the linguist in fieldwork (in the broadest sense). In morphology as opposed to syntax and discourse structure, many data may be systematically elicited in informant work, rather than drawn from texts. See Art. 168.

The object of a grammatical description is a certain language, more precisely, some particular **variety** of a language. The first task of the linguist is to delimit this variety in synchronic and diachronic terms. However narrowly she may delimit it, her empirical data will necessarily be only a sample of the object of description. With a corpus collected in a small village of the Mexican state of Quintana Roo, one may publish a grammar of the speech of those male peasants of that village who were born between 1940 and 1950; or one may publish a grammar of Yucatec Maya. It commonly corresponds both to the purpose of the author and to the interest of the readership to situate the variety described at a rather high level of generality. It is, then, the responsibility of the author to guarantee that her data actually represent her object, i.e. that they are valid. This presupposes some homogeneity of the object. The description of Yucatec Maya based on that sample will be valid to the extent that an observationally similar description could have been produced with data from Valladolid (Yucatan). Mutatis mutandis, the same goes for a grammar of French or Latin.

The actual sources of the examples – informants or texts – have to be identified (the former, if they agree to it). If their representativity could be an issue, it has to be demonstrated that they do represent the object of description. For instance, many a linguistic description is based on a corpus of narrative texts, and it is by no means to be taken for granted that such a corpus does truly represent the language in question. Again, young fieldworkers tend to cooperate with young informants, whose speech usually represents just one sociolect in the community. Therefore, the sociolinguistic coordinates of the data have to be indicated.

It is essential that the data sources be made entirely explicit. The usage observed in traditional grammars of giving examples without indication of the source is no longer acceptable. To put it differently, an example without indication of its source will be

regarded as made up by the author of the description and insofar unreliable (no matter whether she is a native speaker or not).

1.5 Topicality

Scientists are all but resigned to seeing new theories and analyses pop up every day only to be superseded by even newer theories and analyses, so that most of scientific work becomes irrelevant not only to the general public but even to the scientific community itself. Things are different for a grammar (and likewise for a dictionary). First of all, for most languages of the world the publication of a grammar is an event that is not likely to be repeated within a generation. For all those languages that are threatened by extinction, a comprehensive description is the only chance for them to be revived and the only way to protect them from falling into oblivion after their extinction. Secondly, a grammar is a scientific product of potential importance for the general public, since it may and must be the basis of the development of a norm and the composition of primers. All of this heightens the responsibility of the grammarian. The main virtue of the grammarian is (alas) not originality and inspiration but consistency and reliability.

To the extent that grammarians have become aware of these conditions, they have refrained from tying their work to some fashionable model and from concentrating too much on trendy issues. The grammarian has the unenviable task of keeping herself informed on progress in general comparative linguistics but free of ephemeral trends. Fortunately this task has been facilitated in the past decades by comprehensive and up-to-date surveys of general comparative grammar such as Shopen (ed.) 1985 and Payne 1997.

1.6 Concepts and terms

At least two issues arise with respect to the use of concepts and terms in grammatical descriptions: 1) What is the appropriate attitude towards conservatism and neology? 2) What degree of familiarity with concepts and terms should be presupposed?

1.6.1 Conservatism and neology

Originality is a value in science that does not include terminology. Established terminology must be used. Most phenomena in the languages of the world, whether or not they happen to be known to the analyst or to the particular philological or linguistic school that she has been raised in, are actually known in general linguistics and have a term to them which, at least sometimes, is appropriate and well-established. For instance, the semantic relation often called purposive had been called ‘destinative’ by Finno-Ugricists half a century earlier. It is among the duties of the analyst to get informed about established terminology before she coins her own terms. Sometimes, it is true, established terminology is inappropriate. For instance, the term ‘possessive classifier’ was deemed inappropriate even by its coiner, but for want of a better term continues to be used. It is most important that distinct concepts be designated by distinct terms and that these be unequivocal; questions of beauty of terms have less priority.

Secondly, the question regularly arises whether the particular phenomenon in the language being described is simply an instance of something known from other languages.

Here, a middle course between two extremes must be steered (cf. 1.2). One is **particularism**. From the concept of the 'langue' which is both historically and systematically unique, and of the language sign which is arbitrary and unlike anything to be found in other languages, this position deduces that every category of the language must be named by a term of its own in order to avoid misleading identification with something known from other languages. This position has been cherished to the extreme in Russian and American structuralism. It renders a description unusable and blocks comparison of the language with other languages. The other extreme is **universalism**. It assumes a universal grammar and subsumes every phenomenon of the language under pre-established concepts and terms. For instance, the concept of 'case' is stretched to cover (non-morphological) syntactic properties of noun phrases; and then sometimes what used to be called case is renamed 'surface case'. This position is well-represented in generative grammar, but not alien to some brands of functionalism either.

The appropriate middle course is characterized by two propositions: 1) descriptive concepts belong to the level of linguistic types; 2) descriptive concepts are prototypical in nature. Take the concept of 'dative' as an example. For it to be a typological concept means that it is not a universal property of human languages, that some languages have a dative and others don't, that 'dative' is a type of which there are concrete instantiations in the languages and that there are partial functional equivalents to the dative both inside a language that has it and across languages. Being a typological concept for a (value of) a grammatical category, i.e. for a (kind of) linguistic sign, 'dative' is characterized both by its meaning or function (roughly 'case whose basic meaning (Grundbedeutung) is the recipient and whose generic meaning (Gesamtbedeutung) is a participant related indirectly to the situation') and by its structure or form (which for the dative is the same as for 'case', i.e. a grammatical morpheme bound to a nominal expression that marks the latter's syntactic or semantic function in the clause). The prototypicality of such concepts relates to both sides of the linguistic sign. For the meaning or function, it necessitates the Jakobsonian distinction between basic meaning and generic meaning, as just exemplified for 'dative'. For the structure or form, it may involve the identification of a focal instance on a scale of grammaticalization, with some range of variation at both sides. For instance, a case is typically marked by an agglutinative affix, but may be marked by a highly general adposition, on the one hand (as in Japanese or Hawaiian), or by more fusional morphological means, on the other hand (as in Sanskrit).

Of course, such definitions are not always available for grammatical concepts, and often it is the task of the analyst to decide whether an established term covers her particular phenomenon. Extensions of traditional usage are allowed and necessary, because in the last analysis, most of these terms were originally conceived for Greek and Latin and must necessarily be extended in their meaning if they are to be applied to any other language at all. If this were not so, we could never speak of conjugation or of passive formation in English. What is essential here is that such extensions be controlled, i.e. that a general definition (at the typological level) of the kind alluded to be provided, from which it may be deduced that the term is indeed applicable to the particular phenomenon under analysis. If terms are used in this way, then neology in a grammatical description may be kept to a minimum.

The most strongly grammaticalized categories are, at the same time, the most deeply entrenched in a particular language system and consequently the most arbitrary and

language-specific ones. Here the question of whether such a language-specific phenomenon should be subsumed under a category known in typology can become a real issue. An example is provided by the German verb category whose values are traditionally called 'subjunctive I' or 'present subjunctive' (e.g. *singest*) and 'subjunctive II / past subjunctive' (e.g. *sängest*). These terms are actually not helpful as far as the functions of these categories are concerned. If German happened to be an exotic language first described in the second half of the twentieth century, then these categories would probably be called 'hearsay evidential' and 'irrealis', respectively. Consider the two sets of cross-reference markers in Mayan languages as another example. Members of the first set precede the finite verb, members of the second set (displayed in T6 below) follow it. Mayan linguistics, which originates in American structuralism, calls these two paradigms by purely arbitrary labels 'set A' and 'set B', in order to avoid any functional implications. The two sets do combine heterogeneous functions. In Yucatec Maya, set A cross-references the transitive subject and the subject of an intransitive verb in one aspect-mood category, while set B cross-references the subject of a nominal clause, the transitive object and the subject of an intransitive verb in another aspect-mood category. Relating this to certain ergativity-splits, one may say that set A cross-references the subject, while set B cross-references the absolutive. However, this is odd, because a transitive verb would then be flanked by a subject and an absolutive index. Moreover, set A also precedes nouns, cross-referencing the possessor. This function can certainly not be subsumed under the denominator 'subject'. In other words, concepts provided (so far) by typology do not help here; language-specific concepts and terms have to be coined. The only question that remains is whether they must be as empty as 'set A/B'.

1.6.2 Definitions

In principle, the explicit introduction of concepts and terms of general linguistics, especially of analytic concepts such as 'infix' or 'completive aspect', is not the task of a grammatical description, but of genres such as terminological dictionaries and studies in general-comparative grammar. If one could presuppose a systematic methodological organization of linguistic science, a descriptive grammar should only have to categorize its phenomena by the concepts of general-comparative grammar and designate them by the corresponding terms. However, such rigor is not practicable. On the one hand, the extent to which familiarity with established concepts and terms may be presupposed obviously depends on the intended readership of the description. If non-specialists are addressed, technical terminology either has to be introduced explicitly or must not be used at all. On the other hand, established usage may be insufficient with respect to the language being described. Then modifications to it must be made explicit. Still, a grammarian must sometimes be reminded that she is not writing an introduction to linguistics.– Specific suggestions on morphological terminology are found in art. 169, section 3.9.2.

1.7 Description and argumentation

A grammar is not the kind of treatise that has a point. The data are not used to argue for a particular analysis, the analyses are not arranged in a way to convey a certain theoretical insight or to demonstrate a certain method. All of this is necessary and justifiable in other

kinds of scientific treatises. A grammar is, in the first place, a systematic encyclopedia of the grammatical functions and structures of the language.

Consider the French clitic personal pronouns as an example. The grammarian, of course, has a position on the issue of whether they are affixes or clitics. In an article devoted to the subject, it would be appropriate to assemble the evidence for her thesis and argue for it. In a grammar, these pronouns will just be called clitic pronouns (or personal affixes), their various properties will be described, both those like their fixed relative order which are more affix-like and those like their alternative prethematic or postthematic position which are more clitic-like. It is not necessary to turn the description into a persuasory discourse, as is done in some grammars like Cole 1982 and McGregor 1990. Neither is a grammar the place for a linguist to teach linguistic methodology or to demonstrate that she has applied it. What matters is that her analytical decisions be transparent. This is chiefly done by means of examples that exhibit the phenomenon some descriptive statement attributes to the language. It is normally also not necessary to back statements on obligatory rules with ungrammatical examples of constructions that violate them. What is important is that it be made explicit under which conditions the rule applies. True, there will always be cases where the grammarian has not been able to ascertain these conditions. In order not to mislead the reader, she should admit this, for instance by modifying a descriptive statement by the phrase 'under unclear conditions'.

Several grammaticographic treatments (e.g. Comrie & Smith 1977, Noonan to app.) require that the analyst note those properties that the language does not possess. In a comparative grammar, such statements are natural by-products of the comparison. In a monolingual grammar, such statements involve an implicit comparison. Moreover, there is no limit to the things that a language does not have. A feasible form of complying with the requirement is to limit such negative statements to such features that might be expected on genetic, areal or typological grounds and to relate them to the description of some feature the language does possess. Thus, in the section on counting of an onomasiological grammar of German it would be mentioned that the nouns *Stück* 'piece' and *Mann* 'man' can replace the counted noun generically and anaphorically under certain conditions. This may naturally be rounded off by a statement to the effect that this is the closest to numeral classifiers that the language can muster.

2 Grammaticographic problems in morphology

2.1 Tasks of a morphological description

Basically, the core task of the morphological part of a grammar is to describe:

A. **(i)** the internal structure of word forms (grammatical words) (**morphological structure**), and

(ii) the form and the meaning of grammatical items (non-lexical **morphemes**).

These aspects of a morphological description overlap to the extent that 1) the grammatical items of a language are **bound morphemes**, and 2) the internal structure of word forms in this language can be described in terms of these morphemes (typically as a construction of a lexical stem and a chain of bound grammatical items). Since the domain covered by these two conditions is fairly large in many languages, the usual practice of grammarians is not

to separate tasks A(i) and A(ii) from one another for descriptive purposes. As a result, those sub-domains where one of these conditions does not hold commonly come out in the periphery of a morphological description (as represented by sections on, e.g., prepositions or compounding).

Generally, preference is given to what may be termed **word-oriented morphology** (as in A(i)), that is, the word is taken as a starting point for the description. Grammatical items thus fall into two strictly distinct groups which are described quite differently, viz. those which are constituent elements of word structure and those which are words themselves. Within this approach, the very existence of a bound grammatical item constitutes a descriptive statement about word structure. An alternative approach, which may be termed **morpheme-oriented morphology**, would be to start from the morpheme, so that the class of words as well as word classes emerge as a result of the distributional analysis of morphemes. No grammar takes this approach exclusively. However, many grammars contain sections which consist of an annotated list of grammatical formatives. For instance, McGregor (1990), in the relevant sections of his chapter 3, has various lists of grammatical items for each of which he provides the significans, a functional label and a list of functions fulfilled in various constructions. Each item of the latter list contains a reference to other sections where the item is treated in its relevant context.

The word-oriented approach implies a distinction between the morphological description proper (as outlined in A) and two "classificatory" tasks:

B. **(i)** delimiting the class of linguistic units to be referred to as **words** (both in the sense of 'word form' and of 'lexeme'; cf. Art. 26), and

(ii) a classification of lexemes into **parts of speech** (cf. Art. 70) and further grammatically relevant sub-classes, at least inasmuch as morphological patterns vary by word class.

Quite commonly, classificatory solutions are just incorporated into the general structural outline of a grammar, rather than constituting a subject of description in its own right; in other words, the classes of linguistic units dealt with in morphology are presented as established a priori. There are obvious theoretical grounds for this well-established tradition:

First, these tasks are scarcely solved by purely morphological criteria; syntax and semantics play quite a role in an appropriate classification of linguistic units. In this sense, this classification can be regarded as "external" with respect to morphology proper (tasks A). Secondly, this way of presentation overcomes a certain circularity which is apparently inherent in grammatical reasoning, whereby relevant classes are distinguished on the basis of distribution and associated morphological categories, the latter being, in their turn, defined with respect to and in terms of pre-established grammatical classes. However this problem may be solved in the process of grammatical analysis, the usual descriptive practice is to take both the level of the word and the word classes for granted.

On the other hand, classificatory solutions adopted in a given descriptive grammar do constitute autonomous descriptive claims and are even argued for in case they deviate from a "traditional" one, be it a language-specific or the general-linguistic tradition. In addition, minor (sub-)classifications which are relatively or exclusively language-specific (as, e.g., inflectional classes) are more likely to be presented as a subject of morphological description than the major ones. That is, many grammarians take the classificatory tasks to have (at least to some extent) a default solution which can be applied without special

discussion. This default solution is in fact very close to the most traditional parts of speech. By accepting the default solution, a grammarian virtually refrains from a language-specific classification of linguistic units; that is, it is not considered a task of a particular descriptive grammar, but rather a universal ready-made structural template for describing languages.

However, the impact of classificatory solutions on the quality of grammatical description can hardly be overestimated:

1. They play a major role in structuring the morphological description and thus determine the general outline of the grammar (cf. 2.2).
2. They strongly affect the descriptive solutions (tasks A), insofar as **morphological patterns** and **categories** can be defined only with respect to the word classes; once a classification is established, lots of further descriptive solutions are predetermined.
3. They serve as a basis for the **morphology-syntax interface**, inasmuch as syntactic constructions often make reference to a word class and/or a further morphological sub-class of the elements involved.
4. Last not least, these solutions establish the **grammar-lexicon** interface, in the sense that they determine a system of grammatical indices to be provided for items listed in a lexicon (cf. 3.1).

This means that if a language does not fit the default classificatory solution, the consequences of adopting this solution can turn out harmful for the whole enterprise. On the other hand, the default solution is, in a sense, the most user-friendly one, in that it facilitates obtaining information from the grammar, since the reader is faced with a familiar structural outline: the key words (commonly constituting chapter headings, cf. 2.2) tend to be exactly those he expects to find. Thus, a grammarian normally has to weigh the default solution against its descriptive cost in order to find an appropriate compromise. This problem, although often not pronounced as such, seems to be fundamental for descriptive morphology.

A classification into word classes as represented in the morphological part of a grammar is often far more detailed than is needed for morphology proper. The usual descriptive practice is to introduce all more or less grammatically relevant classifications of words by structuring a chapter where the morphological patterns associated with some of these classes are described. In this sense, a description of the internal structure of word forms and a grammar-oriented classification of lexemes are not distinguished in the practice of grammarians, although these – obviously overlapping – descriptive domains presumably never coincide exactly. This practice has the advantage of highlighting the correlation between classifications based on distinct parameters (morphological, syntactic, semantic), the descriptive significance of which is beyond doubt.

To sum up: the practice of descriptive morphology combines three distinct domains of language structure: the word structure, the grammatical items, and the grammar-oriented classification of lexemes, the latter constituting the skeleton of a morphological description. As will be seen below, this combination, however well-motivated, involves some descriptive problems which could be avoided otherwise.

2.2 Structuring a morphological description

The well-established tradition of structuring morphological descriptions is based on a simple and elegant idea of converting grammar-oriented classifications of linguistic units into the section headings of a grammar. This idea makes it possible to combine two descriptive domains (word classes and word structure) in a rather natural fashion, inasmuch as it can be assumed that different word (sub-)classes are associated with distinct morphological patterns; hence, the morphological information is appropriately distributed over these (sub)-classes. The internal structure of chapters varies according to the word class. Roughly, they fall into three groups: open classes with rich morphology, open classes with no or little morphology, and closed classes. A schematic representation of how the respective chapters are most commonly structured is given in T1.

T1. *Types of morphological chapters with respect to their internal structure*

1. **Open word classes with rich morphology:**
   - 1.1. (Optional:) a general overview of morphological structure (e.g., in terms of a general structural template comprising slots for constituent morphemes).
   - 1.2. Further (functional) sub-classification, if relevant for morphology (e.g., transitive vs. intransitive verbs).
   - 1.3. Inflectional paradigm:
     - 1.3.1. The structure of the paradigm(s), i.e., the set of grammatical categories associated with the word class under description and their possible values.
     - 1.3.2. Further (formal) sub-classification into inflection classes (if any) and how the items of the paradigm are constructed for each inflection class.
     - 1.3.3. Exceptions, defective paradigms etc.
     - 1.3.4. (Optional:) semantics and usage of inflectional categories can be picked up as a special issue.
   - 1.4. Word formation (sections organized first by the category of the output of a process, second by the category of the base).
     - 1.4.1. Compounding
     - 1.4.2. Derivation:
       - 1.4.2.1. Regular (productive) derivation; the internal structure is close to a semantically ordered morphemic lexicon (a morphological item + its meaning and distribution).
       - 1.4.2.2. Irregular (non-productive) derivation.
     - 1.4.3. Reduplication, conversion, etc.
   - 1.5. (Optional:) some important semantic and/or derivational sub-classes (e.g., reflexive verbs).
2. **Open word classes with little or no morphology:**
   - 2.1. Semantic sub-classes (e.g., spatial adverbs, temporal adverbs, etc.).
   - 2.2. Correlated structural features (if any) and word formation (often including etymology, grammaticalization notes, etc.).
3. **Closed classes**:
   - 3.1. A general semantic and syntactic overview.
   - 3.2. (Optional:) exhaustive listing of items with their meanings and distribution, ordered according to functional (e.g., interrogative pronouns, indefinite pronouns etc.) or syntactic properties (e.g. postpositions with dative, postpositions with

genitive etc.).

3.2.1. If the items are inflected, the paradigm is given for each.

3.2.2. (Optional:) distinctive structural features of each group (if any) and word formation techniques (e.g., conversion, grammaticalization, etc.).

As shown by this scheme, the morphological parts of grammars contain rather heterogeneous chapters: some are morphological in the word-oriented sense, since they deal with the internal structure of words (cf. the distinction introduced in 2.1); others are morphological in the morpheme-oriented sense, since they describe the form and the functions of closed-class (=grammatical) items; and, finally, there are chapters which are concerned almost exclusively with (non-morphological) sub-classifications of a word class. This may be seen as a direct consequence of the fact that the three descriptive domains combined in morphology (cf. 2.1) remain essentially distinct. Not only do the respective chapters differ in their internal structure, they are bound to take distinct descriptive approaches. In some cases, a semantic classification seems to appear in a grammar just because there is nothing morphological to say about a class. The well-established tradition of providing a semantic classification of adverbs independently of its grammatical significance is a case in point. Also, it seems that if a language has no nominal morphology, the chapter on nouns is much more likely to contain their semantic sub-classification. In fact, a chapter subdivision is the form of introducing a comprehensive classification, but it seems bizarre to leave a section thus generated empty.

A more consequential drawback of the structuring scheme outlined above is that it provides no natural way of accounting for morphological items shared by distinct classes (e.g., for person/number paradigms attached to both verbs and nouns, as, for example, in Yucatec Maya). Another instance of the same problem is represented by grammatical items which function both as a free morpheme (hence, have to be described in a separate chapter of the closed-class type) and as a bound morpheme (hence, a constituent element of some morphological pattern related to a certain word class); cf., e.g., locative prepositions vs. applicative suffixes on verbs in Rwanda (Art. 141, section 3.5.1). Various ways to overcome this drawback of the traditional approach have been applied in different grammatical descriptions. The options are:

1. Introduce **generalized word classes** comprising several word classes sharing some morphological categories (e.g., "nominals"). This solution has a limited domain of applicability because of its hierarchical nature. If the distribution of morphological patterns in a given language does not follow any hierarchical classification (e.g., a sub-class of nouns has an adjectival paradigm, as in Russian), this strategy does not work.
2. Select a “locus” for the description of an item – a morphological category, a single morpheme, or whatever – (for example, the noun for the category of case), and refer to this description in all other relevant chapters (in this case, for instance in the section on pronouns). This solution may work if the semantic and formal properties of the item in question do not vary by word class. If comparable items are even slightly different, a reference will not suffice.
3. Link the descriptions of similar items by cross-references. This is a solution widely applied, obviously necessary, but not sufficient, since it does not provide a way to describe the item with several instantiations as a whole.

4. Separate word formation (for all word classes) from the main body of description and treat it in a special chapter. This compromise solution can overcome the drawbacks of the word-oriented approach at least for derivational morphemes.
5. Separate morphological semantics from the main body of description and treat it in a special chapter. This is a very strong, but rarely applied solution, presumably because of its "ambitious" flavor.
6. Create a **morphemicon** as part of the morphological description as described below (3.5).

The last three structuring solutions tend to be applied in one or another form in recent descriptive grammars. They are, in fact, a concession to the morpheme-oriented approach to grammar. In any case, an appropriate structure for a grammatical description can be achieved only by means of an appropriate combination of various solutions.

2.3 Identification of grammatical items

The task of describing the morphological structure of a word form (cf. A(i)) in 2.2) is almost never addressed directly in descriptive grammars. Instead, it is decomposed into several distinct sub-tasks. The most widely applied decomposition is based on the distinction between **inflection** and **word formation** (in particular, **derivation**; see Art. 38). Its controversial theoretical status notwithstanding, this distinction is widely applied in morphological descriptions. Not only are these types of phenomena often rigidly distinguished (cf. 2.2); these sub-domains of morphology are commonly described in quite disparate fashion:

1. To describe the inflection of a language means to describe each word class in terms of the associated **paradigms**, that is, the inflectional patterns are considered properties of pre-established word classes. To describe the derivation is to describe the semantic impact and distribution of each single **derivational morpheme**; hence, a word sub-class to which some derivational morpheme applies is viewed as a property of that morpheme (this includes the issues of productivity, regularity, etc.).
2. Inflectional forms of different words can be identified as representing the same item of a paradigm independently of their formal similarity; the identity of a paradigm is grasped in terms of certain semantic-functional labels (even if these only imply an approximation of the respective meanings), not in terms of formal identity. In contrast with this, the derivational sub-task involves a semasiological description of particular morphemes, which usually presupposes formal similarity of its various allomorphs.
3. The semantics of inflectional items is assumed by default to be independent of the semantic properties of the word; they are allowed to have a few, usually context-dependent functions. Conversely, a derivational morpheme can have a variety of meanings depending on individual stems it is applied to.

These properties can be traced back to the concepts of "prototypical" inflection and derivation and the role they are assumed to play in the grammar and, in particular, in the grammar-lexicon interface (Art. 36). The distinction between inflection and derivation thus determines the choice of a descriptive perspective. That is, there is a considerable difference in how inflectional and derivational items are treated in descriptive grammars, from the identity of an item up to the types of information to be provided.

In particular, the concept of inflection implies a more function-oriented identification procedure, that is, both formal identity of items even in case of clearly related meanings, and an obvious formal contrast are very easily disregarded, as far as the inflection is concerned (hence, multiple non-distinctions within paradigms; see Plank (1992) for an overview of this problem, and Zaliznyak (1967: 19-34, 129-148) for case paradigms). Here is an example: Russian is an aspect-dominated language in the sense that the grammatical category of tense works differently for perfective and imperfective verbs. Tense allows for two different ways of description, schematically represented in T2 by the tense forms of the verb pair *delat'* 'make/do (impfv.)' and *sdelat'* 'make/do (prfv.)':

T2. *Russian tense: alternate paradigms*

1. form-based description		**2. function-based description**		
			imperfective verbs	**perfective verbs**
+Past	*dela-**L**, sdela-**L***	**Past**	*dela-**L***	*sdela-**L***
–Past	*dela-**ET**, sdela-**ET***	**Present**	*dela-**ET***	
Imperf. Future	*BUDET delat'*	**Future**	***BUDET** delat'*	*sdela-**ET***

Basically, the two accounts differ in how they treat forms like *sdela-**ET*** (PFV:make-FUT/NPAST.3SG) 'will make, will have made'. Variant 1 highlights the formal identity between this item and the Imperfective Non-Past *dela-**ET***. They are identified as instantiating the same item of the tense paradigm and get the same functional label (Non-Past), see Comrie (1978: 66-67). Yet these forms are clearly distinct semantically: the form of perfective verbs can refer only to the future, while with imperfective verbs it is employed basically for reference to the present (although the future meaning is not excluded). Thus, the descriptive cost of this solution is a dependency of the range of functions of an inflectional item on the verb class.

By contrast, variant 2 disregards the formal identity between the items under discussion. Instead, it highlights their functional distinction and, accordingly, the functional similarity between the Perfective Non-Past (Future) of the perfective verbs and the Imperfective (analytical) Future. The latter pair of forms are assigned the same position in the paradigm and get the same functional label (Future), despite the obvious formal contrast. The dependency on the aspectual meaning of the verb stem is thereby built into the structure of the paradigm, instead of constituting an independent descriptive statement. The descriptive cost of this solution is (i) the formal heterogeneity of the Future and (ii) the unmotivated formal identity of the Perfective Future and the Imperfective Present (which would show up recurrently in the description of multiple formal variants of this marking, only one of which is represented in T2).

Variant 2 is apparently preferred in existent descriptions of Russian (even though some of them do suggest variant 1, most often the formal identity between the Imperfective Present and the Perfective Future is not mentioned). In other words, the paradigm is defined in such a way as to simplify the semantic description at the cost of somewhat artificial formal complexity.

Formal identity of this type is never neglected as easily for derivational formatives; to continue with Russian examples, its verbal postfix *-s'a/-s'* expresses an extremely wide variety of meanings in combition with different verb stems (middle, reflexive, reciprocal, anticausative, dispersive, to mention only some of them). Yet these are commonly treated as instances of the same linguistic unit.

In view of this difference in descriptive perspective, the well-known problem of delimiting inflection and derivation gets a new kind of relevance: how a grammatical item is identified and described depends on its affiliation with inflection vs. derivation, at least above a certain degree of discrepancy between form and meaning. It might be the case that these two sub-domains of morphology are in fact delimited in descriptive grammars depending on which approach seems more "suitable" for a given item. On the other hand, the choice of one or another perspective can affect the very properties of the items identified (as they would be described in the grammar), inasmuch as the prototypical features of inflection vs. derivation are embodied in the identification procedure.

3 Structure of a grammar

3.1 Comprehensive presentation of a language

No grammar can be complete. Given limitations of every kind, the author has to set priorities. These follow from the main purpose of the grammar and from external conditions. Needless to say, on the basis of available publications, one may decide to produce a partial description. However, the aim may be to produce a comprehensive presentation of the language. Assume a field structured in a hierarchical fashion. For an account of it to be comprehensive means that it is balanced in terms of the amount of detail provided for each of the sections at a given level of the hierarchy. In this sense, a grammar such as McQuown 1990 (of Totonac), which comprises 64 pages on phonology, 107 pages on morphology and 11 pages on syntax is unbalanced and therefore not comprehensive.

A comprehensive account of a language is articulated on the three levels shown in T3.

T3. *Levels of presentation of a language*

1 the **documentation** of the language, which is a corpus representing (analyzed) primary data;
2 the **description** of the language, whose object are the data of level 1;
3 a **methodological reflection** on the description, whose object is the account of level 2.

The levels are, thus, in a meta-relation to each other. The reflection of level 3 has the function of accounting for the purpose of the description, the conditions under which it was carried out, including the achievements of previous scholarship, its theoretical and methodological prerequisites, the many choices and decisions that the author has made; and as mentioned in 1.4, it reflects on the character and limitations of the data base. Most of this part commonly takes the form of an introduction to the description. Some of it may be relegated to an appendix.

The corpus on which the grammar is based (level 1) will normally not be reproduced in full. Good modern presentations of a language contain an appendix that presents some specimina of representative texts (see Mosel 2002, §6.2 for the scientific importance of such a collection). These are provided in the form of video or audio recordings and rendered in the canonical trilinear representation as explained in Art. 169. More on the relationship between the corpus and examples in the running description in 4.5.

The descriptive part of a comprehensive account of a language (level 2) consists of two subparts, the system of the language and the setting of the language. Here we may be brief on the latter. The **setting of the language** comprises an explanation of relevant glossonyms, the genetic affiliation of the language and its dialects, its ethnographic situation (i.e. the situation of its speech-community), its cultural situation (including, importantly, its written tradition) and its sociolinguistic situation, i.e. its internal stratification, its status in the speech community and its areal relations. This part of the description provides a referential background for parametricization by diachronic, diatopic, diastratic and diaphasic variables that will prove necessary in the account of the language system. It is, at the same time, the one part of a comprehensive presentation of a language that is most likely to take a historical perspective even if the rest is purely synchronic.

The **language system** has two main parts, the expressive subsystems and the significative subsystems (corresponding to Martinet's second and first articulation, respectively). The primary expressive subsystem is the phonology (with the phonetics), the secondary one is normally the writing system. (This is, of course, also true for [ancient] corpus languages.) The significative subsystems are not alternative, but jointly exhaustive: the grammar and the lexicon. They are significative because their units embody a mapping of meanings onto expressions. The significative subsystems are articulated in terms of levels of complexity: At the level of the word form and below, we have morphology, which – in the form of inflection – is an essential part of the grammar and also – in the form of word formation – of the lexicon (see Art. 36). At higher levels, we have the syntax as part of the grammar, and phraseology as its counterpart inside the lexicon.

Since there is, in principle, no borderline between grammar_1 and lexicon, there is also overlap between a grammar_2 and a dictionary of a language. The morphology describes the word-formation patterns of the language. Their products are nevertheless listed in the dictionary. Only in the limiting case of a word-formation process that applies completely regularly and productively – e.g. formation of certain verbal nouns – may one renounce to representing each of the products in a lexical entry of its own. Contrariwise, the dictionary will contain a couple of entries that result from a word-formation process that is no longer productive in the language. Mentioning it in the morphology would entail repeating the same list of items that is in the dictionary. In such cases, redundancy is preferred to parsimony. The drawbacks of redundancy are that it induces inconsistency and that it may be uneconomical. They are outweighed by the advantages, which include user-friendliness and theoretical soundness (the mental grammar and lexicon are redundant in the same way). Moreover, we are not talking about a literal repetition of material in two parts of the description, since the ordering principles and, consequently, the ways of accessing the information are entirely different between dictionary and grammar_2.

Another important relationship between the grammar and the dictionary is that the former introduces and defines the terms that appear in the cells on grammatical information of the microstructure of a dictionary entry. This concerns such categories as noun class, gender, countability, aktionsart, inflection class etc. Technically speaking, this kind of specification in a dictionary entry determines whether the item can be used in a certain construction described in the grammar.

3.2 Onomasiology and semasiology

As we said in section 1.2, the fundamental problem of grammaticography is to provide a common format for descriptive grammars while at the same time taking care not to obliterate the individuality of the language being described by forcing it into a Procrustean bed. The general task of a language is to provide a mapping between meanings and expressions. The meanings have an extra-linguistic substrate (i.e. one independent either of language altogether or at least of particular languages) in cognition and social interaction (communication). The expressions have an extra-linguistic substrate in phonetics and semiotics. It is the mapping itself, achieved in the grammar and in the lexicon, that is proper to each language. There can therefore be a universal system of cognitive and communicative domains coded by languages and a universal system of expression techniques and sounds used by languages. There cannot be a universal system of grammar.

A system of lexical and grammatical description that provides comparability of the object language with other languages can therefore be either based on a system of cognitive and communicative domains and then describe how the language in question manifests each facet of these in its expressions; or else it can be based on a system of expressive (structural) devices and then describe what each of them is used for in the object language. The former approach is traditionally called **onomasiological** (or synthetic, more recently 'functional'), the latter **semasiological** (or analytic, more recently 'structural') (cf. first Gabelentz 1891/1901: 84-104 and, more recently, Lehmann 1980, 2002[P], ch. 1.2.1 and Mosel 2003, §7). Each of these approaches is in itself coherent and capable of providing a complete description. Many grammars stick to one of them. For instance, Jespersen 1937 is a purely semasiological grammar, while Givón 1993 is a purely onomasiological grammar. However, each of these approaches is one-sided, as the onomasiological approach corresponds to the viewpoint of the speaker, while the semasiological approach corresponds to the viewpoint of the hearer. Thus, an onomasiological grammar answers questions of the type: how can I express such and such a thought, or fulfill such and such a communicative function, in this language?, while a semasiological grammar answers questions of the type: what does such and such an expression of this language mean? Since grammars, just as dictionaries, are generally meant to serve both the speaker and the hearer, the ideal grammar consists of two parts, an onomasiological and a semasiological one. It should be clear that such an arrangement is also maximally user-friendly, because whatever question a user may pose to a grammar is posed either in the speaker or in the hearer perspective. Moreover, most of the descriptive problems reviewed in 2 resolve themselves if the twofold approach is chosen.

A couple of published grammars approach this ideal to some extent. Despite appearances and declarations of the author, Gabelentz 1881, a grammar of Classical Chinese, is relatively far off his own mark because the "synthetic grammar" is not really onomasiological, but just a construction-based syntax. McGregor 1990, a grammar of Gooniyandi, organizes its chapters in the familiar bottom-up fashion appropriate for a semasiological account. However, some syntactic chapters and part of a final chapter on semantics take an onomasiological perspective on the items introduced in lower-level sections. The chapter on morphology in Haspelmath 1993, a grammar of Lezgian, contains sections on noun morphology and verb morphology, each of which is subdivided into a form-based and a function-based subsection; but otherwise the grammar mixes the two

approaches. Lehmann 2002[P] is not a grammar, but just a description of possession in Yucatec Maya. It is subdivided into a chapter that introduces the relevant structures in a semasiological perspective, and three chapters that take the opposite perspective. It must be said that to this day, most grammars mix the two approaches in uncontrolled ways (cf. Lehmann 2004[F]).

It is nowadays standard to tie the distinction between morphology and syntax to the distinction between levels of grammatical complexity. However, it has often been associated with the distinction between an onomasiological and a semasiological grammar. From traditional grammars of the nineteenth century to our day, many a grammarian has said that her morphology deals with the system and structure of linguistic forms, while her syntax deals with their use. It might thus appear that the morphology of such a grammar is semasiological, while the syntax is onomasiological. However, such systematicity is only apparent. Most traditional morphologies contain chapters on functional categories such as person and tense – and to this extent they are onomasiological. And the typical traditional syntax has a chapter on the use of the cases, thus starting off from a structural concept introduced in the morphology and tracing its functions – and to this extent it is semasiological.

Actually, the canonical level-dependent distinction between morphology and syntax can only be made in a semasiological grammar. A semasiological grammar is organized according to the hierarchy of structural complexity of linguistic units. For all those languages which possess the word form as a level of grammatical structure, the section of semasiological grammar devoted to this and lower levels will be the morphology. In the last quarter of the twentieth century, there have also been functional (i.e. onomasiological) grammars that presupposed a distinction between morphology and syntax; e.g. Comrie & Smith 1977 and the series of grammars based on it. Since this distinction cannot be made in an onomasiological grammar, it introduces inconsistency into it. For instance, the grammar of many languages manifests the concept of definiteness, but some do so at the level of morphology, while others do so at the level of syntax.

In the twentieth century, many structural (thus, semasiological) grammars were published. Contrariwise, there are as yet few purely functional grammars. A theory of cognitive and communicative domains of language started to be developed only in the last quarter of the twentieth century, and for most of these domains we still lack both a solid foundation and an internal structure adequate for linguistic description. The present handbook tries to partly compensate for this shortcoming by organizing those of its chapters that are devoted to grammatical categories and operations (XII and XIV) according to an onomasiological perspective. It is inevitable that the articles of those chapters break the boundaries between morphology and syntax. The handbook on syntax in this series might have done the same with equal or greater justification.

3.3 Structure of a semasiological grammar

Much of the language system is structured in terms of hierarchies of levels of complexity. This concerns, first of all, the system of grammar$_1$ as mentioned in 3.1; but it may also include, with some imprecision, the distinction between expressive and significative systems. This yields the well-known series ‘phonology – morphology – syntax – discourse’, and inside grammar the sequence ‘stem – word form – phrase/syntagm – clause

– (complex) sentence'. The history of linguistics has, *cum grano salis*, followed this progression. Syntax was a step-child of linguistic description up to the end of the nineteenth century, and even the 20th century has seen many "grammars" that actually boil down to a phonology plus a morphology. Discourse has come to be studied systematically only in the last third of the 20th century. The progression is deeply entrenched in the awareness of linguists, who think it must shape the organization of their grammar. Almost all grammars work essentially bottom-up (cf. Mosel 2003, §5.2). This seems to correspond to a didactic progression which starts from elementary units and proceeds stepwise to complex units. A top-down progression, as it is prescribed in Comrie & Smith 1977 and followed in *Lingua Descriptive Studies*, seems unnatural because in the treatment of a given unit (e.g. the complex sentence) one is forced to appeal to constituent concepts (the clause, in this case) which have not yet been introduced.

However, the picture must be modified slightly. The bottom-up approach corresponds to analytical thinking, which combines elementary units according to rules and aims at constructing a complex whole in a compositional fashion. The top-down approach corresponds to holistic thinking, which starts from a whole and understands its parts in terms of their function in the whole. As is well-known, the two approaches do not exclude, but complement each other. To give an example: A cleft-sentence is best understood if one knows what contrastive focus is and what it entails for semanto-syntactic structure. One can then identify the structural constituents in terms of their role in the complex construction. Contrariwise, it will be hard to construct the purport of a cleft-sentence in a bottom-up fashion, starting from a copular predication over an empty subject that combines with something that looks like a relative clause. More precisely, although the chapter on the cleft-sentence does presuppose the notions of copula clause, relative clause and complement clause, it is not the case that the grammar can build on one of these and expand it into a cleft-sentence. The external grammatical relations of a given unit – the subordinate clause in this example – are not treated in the chapter dealing with this unit, but instead in the chapter of a higher unit whose internal relations they are and of which the given unit is a constituent. This responds to the principle – last put forth in construction grammar – that the formation of complex constructions is goal-directed.

The same goes for the morphology. The semasiological description of word-formation does not start from a certain derivational suffix, combine it with bases of different categories and then look what the category of the result is. Instead, there is a section on stems of a certain category, e.g. the adjective. The category has certain elementary members, i.e. adjectival roots. Next there are possibilities of forming adjective stems by various formal processes, e.g. by suffixal derivation. One of them is our derivational suffix. Finally, it is seen that it may combine with bases of various categories to yield the result at stake.

The general principles of such an arrangement of a semasiological description may be formulated as follows.

1. The description works bottom-up through the **hierarchy of grammatical levels**.

2. For the grammatical unit of each of these levels, the categories into which it is articulated are identified. For each of the potentially complex categories, its internal syntagmatic structure is analyzed: First, a set of constructions according to kinds of syntagmatic relation is enumerated. For each of these constructions, the nature and distribution of its elements is set forth. Finally, given a certain construction of elements of

two categories, one of the categories may comprise a grammatical (in particular morphological) paradigm of elements. Such paradigms are discussed as part of the description of the particular construction.

3. Only the internal syntagmatic structure of a given unit is part of the treatment of that unit. Any structural phenomena which concern the relation of a given element to its context are treated at the point where the including construction – the one which provides the context – is treated.

To give a final example: At the level of the word form, categories such as 'finite verb form' are identified. There are kinds of finite verb forms according to their internal structure, e.g. periphrastic forms and synthetic forms of different kinds. One of these is the synthetic form that consists of a tensed stem and a personal ending. In the latter position, there is a paradigm of morphemes whose internal structure is treated now. It may be seen that this approach combines a bottom-up progression to ever more complex constructions with a top-down analysis of each of these constructions.

3.4 Structure of an onomasiological grammar

The cognitive and communicative domains that are coded in language comprise concepts and operations a subset of which manifest themselves in grammatical structure. These are such concepts as the addressee or directed motion and such operations as abstraction or making a question. Such concepts and operations are assembled in **functional domains**. They are by now well-researched in functional typology, so that they can be tentatively enumerated. Since they provide the highest-level subdivision of an onomasiological grammar, this will be done here in T4.

T4. *Functional domains*

domain	**basic functions**	**representative concepts and operations**
apprehension & nomination	an entity is grasped by categorizing and individuating it; it is named by a label or a descriptive expression	categorization, types of concepts, empathy
concept modification	a concept is enriched, or an object is identified	attribution, apposition, relativization
reference	a representation is related to and delimited within the universe of discourse	determination, deixis, reference tracking
possession	the relation of an entity to another one is established or inheres in one of them	possession in reference, possessive predication, external possessors
spatial orientation	an entity is localized in space statically or dynamically	reference points, local relations, spatial and gestalt properties of objects
quantification	the extent of the involvement of a set of entities in a predication is delimited	quantification in reference and in predication; counting, ordering
predication	information is attributed to a referent	existence, situation, characterization
participation	a situation is articulated into an immaterial center and a set of participants and circumstants related to it and to each other	control & affectedness, central vs. peripheral roles, alignment of fundamental relations
temporal orientation	a situation is designed with respect to its internal temporal structure and limits and temporally related to another situation	situation types, aspectuality, temporal relations
illocution, modality, evidentiality	a proposition is rendered relative to speaker, hearer and reality	speech acts, obligation, volition, possibility, toning, evidentiality
contrast	a concept or proposition is assessed qualitatively by comparison with similar ones	negation, comparison, gradation, intensification
nexion	a situation is expanded into a complex one, or several situations are linked together	speech reproduction, complementation, interpropositional relations
communicative	a proposition is articulated in foreground and background	discourse structure, functional sentence perspective

dynamism		(topicalization, focusing, emphasis)

Onomasiological description proceeds from very general cognitive and communicative functions as indicated in the central column of T4 through more specific subareas such as those of the right-hand column down to the functional categories and processes of the language under description. These are finally mapped onto the structural devices and grammatical formatives introduced in the semasiology. For instance, in a grammar of German, the chapter on possession treats, among other things, possessive pronouns, the genitive attribute and its equivalents, the possessive dative, possessive predications with *haben* and *sein* etc. Most of these structural devices recur also in other chapters of the onomasiological description; but this chapter is where the question is answered how German expresses an attribution of possession as in Turkish *vakt-im var* (time-POSS.1.SG EXIST) 'I have time'.

3.5 Additional parts of a grammar

The subdivision of a grammar follows mainly from the overall organization of the presentation of a language as outlined above. Since the term 'grammar' is not seldom used in a broad sense almost equivalent to 'language description', it bears repeating that, conceptually, the data corpus, on the one hand, and the methodological reflection, on the other hand, are not part of the grammar; and neither are the account of the setting of the language, the lexicon, the phonetics, phonology and orthography. What remains is simply the morphology and the syntax. Of the morphology, inflection is properly included in grammar, while word formation could, in principle, be treated either in the dictionary or in the grammar. For practical reasons, it is always included in the grammar (if it is treated at all).

An important challenge taken up in many modern grammars written on a typological background is to present an epitome of what the language is like and how it works, i.e. to present the language as instantiating a **linguistic type**. From what has been said in 1.7 it follows that the overall presentation style of a reference grammar does not by itself fulfill this task. It is therefore widespread practice to provide an additional chapter – usually at the beginning of the grammar – that characterizes the language at the typological level. While the main body of a grammar is descriptive, this chapter provides the opportunity to compare the language with other languages and to answer the Humboldtian (1836:417) question "auf welche Art [sie] die hauptsächlichen Fragen löst, welche aller Spracherzeugung als Aufgaben vorliegen."

Furthermore, a description of a language system contains a number of lists. The most important of these is, of course, the list of significative units (signs) of the language. It comprises the lexemes (including phraseologisms) in their citation form plus the morphemes, i.e. the roots and the grammatical and derivational formatives, including the affixes. The question is in which part of the overall description and in which way these lists are presented.

In the ideal situation, there is a **dictionary** of the language beside the grammar. If so, then the dictionary contains entries for all of these items. It then properly includes a **morphemicon** (inventory of morphemes), i.e. morphemes are lemmata just like stems or citation forms of words. For the sake of user-friendliness, not only morphemes, but also variants are conceded an entry, the latter reducing to a reference to the main entry. A dictionary entry refers to the grammar in two ways: 1) implicitly by the grammatical

categorization and other grammatical information provided in the microstructure of each entry (see 3.1); 2) explicitly in the form of references to the relevant sections of the grammar for such entries which are grammatical formatives. In such a situation, the grammar contains no morphemicon. Instead, each grammatical morpheme is introduced in the semasiological description as part of its construction and paradigm. The distribution of information on individual grammatical items between dictionary and grammar is then a question of fine-tuning. There must be a certain amount of overlap as regards generic properties. Details on meaning and function, especially idiosyncratic properties, are provided in the lexicon, details on distribution and conditions of variation are provided in the grammar.

If there is no extra dictionary, then a number of second-best solutions are available. Of these, the relative best is a **glossary** or vocabulary appended to the grammar that contains all the lexemes and morphemes that are mentioned in the grammar, including the examples, each coupled with its meaning or function. Again, variants are listed, too; and for the grammatical items there are references to all the places where they are discussed. The absolute minimum requirement is an alphabetical **index of grammatical items** at the end of the work, again with the relevant references.

Among the indices there is also a **subject index**. While this is nothing special of books on grammar, it does assume a special role here. If the grammar is bipartite and both the semasiology and the onomasiology are organized hierarchically according to established conceptions, then the user can retrieve the information by systematic search through the table of contents if this is sufficiently detailed. Even then the grammar probably needs an index, as there are liable to be at least some concepts and terms that are language-specific or appear in unwonted contexts. But as long as grammars are not organized in such a way, an index of functional terms is necessary in a structural grammar to compensate for the lack of the opposite perspective; and vice versa for a functional grammar.

Another kind of list that is often attached to a grammar – and non seldom also to a dictionary – is the list of inflection paradigms (mostly of the specimen type; see 4.3) of various word classes and their subclasses. If the language has a relatively complicated inflectional morphology, as e.g. Ancient Greek conjugation, then it makes sense not to interrupt its systematic description by pages filled with conjugation tables and instead to relegate these to an appendix.

Among the various lists that a grammar shares with other genres, such as lists of tables, figures, primary text sources, bibliographical references and the like, the list of **abbreviations** is of special interest. Among all the abbreviations used in a grammar, the (abbreviations of) grammatical category labels form a distinct subset. Whether or not other kinds of abbreviations are listed, the latter must certainly be listed, and they may constitute a list of its own. A separate list of grammatical category labels also helps the author to be consistent.

4 Descriptive devices

4.1 Representation of a single grammatical item

4.1.1 Representation of the significans

Basically, one representation type is chosen for the whole grammatical description (i.e. excepting the phonetics and phonology). The primary choice is between a standard **orthographic representation** and a (technical) linguistic representation. The former is generally preferred because it makes the description more accessible to the speech community and other non-linguists interested in the language. Sometimes, standard orthography is supplemented by additional diacritics, e.g. to mark word stress.

There are, however, situations where a linguistic representation is necessary or preferable, if a language has no orthography or if using traditional orthography would complicate comprehension of the grammar and is not in accordance with its descriptive tasks (cf., e.g., recent Chinese grammars). In this case, authors choose a basic representation trying to obtain a compromise between its theoretical ambitions, convenience for descriptive purposes and user-friendliness (especially in a learner's grammar, the latter includes easy rules of reading). Sometimes printing (publishing) facilities prove to be a limiting factor. If the representation is based on linguistic principles, it is either a **broad phonetic transcription** or a **phonemic** or a **morphophonemic representation**. Use of a phonetic transcription as the basic representation in a grammar is very rare and essentially limited to situations in which a grammar is published before a satisfactory phonological analysis of the language is available. In general, a morphophonemic representation is to be preferred to a phonemic one, but lower-level representations are, of course, necessary in those sections that deal with lower level variation. In phonetic representations, symbols are chosen from the International Phonetic Alphabet. At higher levels and if a certain phonetic distinction is irrelevant in the language, usually less technical letter symbols are preferred.

The basic representation is used throughout a grammar; yet, a description of morphology often involves more or less significant deviations from this representation. In particular, an orthographic representation can prove insufficient for an adequate and sufficiently detailed morphemic segmentation, either in a language with intricate morphonology or in case the traditional orthography obscures certain morphological phenomena. Hence, some additional elements can be incorporated into representations of some forms or even phrasal examples in order to provide a transparent morphemic segmentation.

For example, Russian orthography makes use of graphemes like *я*, *ю* which stand for combinations /ja/, /ju/ after a vowel or for single phonemes /a/, /u/ after a palatal consonant (very loosely, in the latter environment these graphemes are employed to signify the palatalization of the preceding consonant, in contrast with their plain equivalents *а, у* which occur after non-palatal consonants). This spelling convention makes the orthographic representation unsuitable for morphemic segmentation and for the representation of single morphemes, as is necessary, i.a., in combination with interlinear morphemic glossing (Art. 169). In a strictly orthography-based Russian grammar one would find numerous items like an "ending *-а/-я*", a representation which suggests allomorphy, but in fact just reflects the fact that the ending /a/ occurs both after non-palatal

and palatal consonants (in particular, after /j/). In order to avoid such misleading representations, a spelling convention can be employed which renders /j/ between vowels as well as palatal consonants in a more transparent fashion (Zaliznyak 1967: 201-294). As an example, consider the following forms, where morphemic borders are indicated as imposed by the orthographic representation:

	student	**father-in-law**	**hero**
sg.nom	*студент*	*тесть*	*герой*
sg.gen	*студент-а*	*тест-я*	*геро-я*
sg.dat	*студент-у*	*тест-ю*	*геро-ю*

The phonemic representation is:

	student	**father-in-law**	**hero**
sg.nom	stud'ent	t'es't'	g'iroj
sg.gen	stud'ent-a	t'es't'-a	g'iroj-a
sg.dat	stud'ent-u	t'es't'-u	g'iroj-u

It is now clear that the endings do not vary between these words; yet the transcription cannot help but contain deviations from the orthography which are of no relevance here and can only be misleading (cf., e.g., /i/ for *е* in the last column). Now consider the same forms in an alternative spelling (the deviations from the established convention are shown by capital letters):

	student	**father-in-law**	**hero**
sg.nom	*студент*	*тесть*	*герой*
sg.gen	*студент-а*	*тестЬ-а*	*геройЙ-А*
sg.dat	*студент-у*	*тестЬ-у*	*геройЙ-У*

The orthographic representation is brought closer to the phonemic one in order to show that the three nouns follow the same inflection pattern, yet the deviations are reduced to a necessary minimum. The symbols reflecting a distinct representation type are underscored either by capitalizing or by using symbols from another alphabet (continuing the Russian example above, the symbols /'/ and /j/ might be used instead of *Ь* and *Й*). Such mixed representations may be used as a compromise between the incompatible requirements of using a standard orthography and providing an adequate morphological analysis.

Another side of the same problem is the representation of a morpheme abstracted from its morphological environments; that is, technically speaking, a unitary representation of a set of allomorphs. Several techniques (which may also be combined) are in use:

1. Regular phonemic alternations are ignored (a technique normally favored by the use of the orthographic representation).
2. For more or less regular morphonemic phenomena, special symbols for morphonemes are employed, commonly, a capitalized phoneme symbol (for instance to abstract from vowel harmony phenomena).
3. The allomorphs are just listed, commonly separated by a slash.

There seems to be no established means to represent the significans of non-segmental items (e.g. metaphony).

4.1.2 Representation of the significatum

The meaning of a grammatical morpheme may be rendered in several ways, among them:

- by using an appropriate functional (mnemonic) term, e.g. ‘iterative’, ‘ablative case’, etc., which may then be referred to by means of a conventional abbreviation, e.g. ITER, ABL etc., where upper case identifies (abbreviations of) grammatical category labels;
- by translating it into the metalanguage (e.g. ‘constantly’, 'from').

The choice of a name for a category may be determined by a language-specific terminological tradition or by a desire to keep the terminology as transparent and “speaking” as possible; see 1.6.1.

Semantic information in morphology almost never reduces to naming categories appropriately (although, in some cases, a reference to the syntactic part of grammar is considered sufficient). The semantics of derivational morphemes is commonly described in terms of processes and operations. This is appropriate wherever operations apply in a regular and productive way, yet it can be misleading in the case of fossilized patterns. In the latter case, restrictions on the distribution should be carefully described (sometimes, just a list of relevant items seems the best solution, to be sure, with comments on their semantic motivation). The semantics of inflectional items is commonly described in terms of grammatical features structuring the respective paradigm; cf. 4.2.

To visualize the range of meaning of a polysemous (multifunctional) morpheme, a **semantic map** is a useful descriptive device (see Haspelmath 2003). This is a language-independent n-dimensional (mostly with n=2) arrangement of monosemic grammatical functions by their similarity. A semantic map of a particular case, e.g. the Latin dative, will cover a contiguous area that comprises a subset of case functions (recipient, addressee, experiencer etc.) appearing on the map. A similar map can be drawn for the dative of another language, e.g. Turkish, and then the semantic expansion of the dative in the two languages may be easily compared. Such a semantic space is also the locus for diachronic semantic variation.

4.2 Category and feature

Rules of syntax may refer to a morphological category or **grammatical feature** that a constituent bears. The formulation of such rules often presupposes the specification of such categories in a parameter-value format, where a grammatical category constitutes a parameter, and its subcategories, the values (s. Art. 28). Thus, the inflectional information of a Latin verb form may be represented as follows:

cantaveritis

2	person
m	number
u	tense
m	anteriority
m	mood
u	voice

Another task of a morphological description is to provide an appropriate system of **inflection classes** to be picked up in a lexicon (cf. 2.2 and Art. 65). The following duties await the analyst here:

1. The inflection classes have to be brought into a hierarchy. Wherever a language has some complexity in this area, major inflection classes usually have minor subclasses, and these have a few totally irregular members. For instance, one of the major conjugation classes in Latin is the consonantal conjugation. One of its subclasses is constituted by a perfect formation which involves lengthening of the root vowel. This, in turn, has *emō* 'buy' as a member, which is slightly irregular in having an epenthetic consonant in the perfect participle *emptum* 'bought'.
2. The inflection classes have to be named by appropriate terms. Since the conditioning factor of an inflection class is often the stem final, this is suitable as a label of the class, as in the Latin *a* conjugation. The concept of a particular inflection class reflects its position in the taxonomy, e.g. 'consonantal conjugation with lengthening perfect'. This, of course, leads to cumbersome terms which may be abbreviated by numbers and letters, e.g. '3e' (third conjugation, subclass e). The abbreviations are needed, among other things, in the section of grammatical information of the lexical entry of such a verb.
3. For didactic purposes, those inflected forms of the paradigm are identified from which one can deduce the entire inflectional paradigm. For this reason, Latin nouns are often quoted not only in the nominative, but also the genitive, e.g. *gēns, gentis* 'stem'. For German, one even needs the nominative plural in addition (*Hahn, Hahns, Hähne* 'rooster'). For many semiregular Latin verbs, if one knows the first person singular present indicative active, the first person singular perfect indicative active and the perfect participle, then one can conjugate the verb correctly through all of its categories. Traditional grammars and dictionaries therefore specify the forms of these categories for semiregular verbs, e.g. *emō – ēmi – emptum* 'buy'. Instead of such a set of word forms, it is also possible to mention a set of allomorphs appearing in such word forms. For instance, a German apophony class may be identified by *i – a – u* (as in *singe – sang – gesungen*).

4.3 Paradigms

Inflectional morphology is commonly represented in terms of paradigms arranged as one- or multidimensional tables. For each word class, the set of tables has the same structure in terms of number of columns and rows and their category labels. There are two linguistic forms of such tabular representations. In the traditional paradigm, a plain example word is represented in the forms of all the values of the categories, as in T5.

T5. *Specimen paradigm of Latin personal endings*

number / **person**	**singular**	**plural**
1	laud-o	laud-amus
2	laud-as	laud-atis
3	laud-at	laud-ant

This kind of representation is the original linguistic sense of the Greek term *parádeigma* 'example'. It is still used in textbooks or if the morphology is so fusional that inflectional markers are hard to represent separately.

This leads us to the second form of representing a paradigm, in which the cells of the table are occupied just by inflectional morphemes, as in T6.

T6. *Marker paradigm of Yucatec personal suffixes*

number / person	singular	plural
1	-en	-o'n
2	-ech	-e'x
3	–	-o'b

This kind of representation is more abstract and proper to linguistic science. It is viable for agglutinative morphology. The **specimen paradigm** is, in a sense, a definition *per ostensionem* that appeals to intuitive understanding. The **marker paradigm** only works to the extent that the information that it presupposes is made explicit in the grammar. This concerns, in particular, syntagmatic information on the way the markers combine with the stem (see 4.4). If complex morphophonemics are involved, it is prudent and user-friendly to complement the rules by a specimen paradigm.

For a given morphological category, the order in which its values are enumerated is fixed by tradition in many cases; e.g. '1st, 2nd, 3rd person', 'nominative, genitive, dative, accusative, ablative, vocative case'. A properly linguistic criterion for their arrangement is **syncretism** (Art. 66). Especially for semantically motivated syncretism, it is convenient if the syncretistic category values are adjacent in the table. T7 and T8 are two examples to illustrate ways of displaying syncretistic paradigms:

T7. *Declension of Old Indic a-stems (paradigm of deva- 'god')*

number / case	Singular	Plural	Dual
vocative	*deva*	*devās*	*devau*
nominative	*devas*		
accusative	devam	devān	
instrumental	devena	devais	devābhyām
dative	devāya	devebhyas	
ablative	devāt		
genitive	devasya	devānām	devayos
locative	deve	deveṣu	

T8. *Russian conjugation*

	tense	**present**			**past**		
number	**gender / person**	**m**	**f**	**n**	**m**	**f**	**n**
sg.	**1**	čitáju			čitál	čitála	čitálo

	2	čitáeš			
	3	čitáet			
pl.	**1**	čitáem	čitáli		
	2	čitáete			
	3	čitájut			

The tables representing paradigms are at most two-dimensional. For agglutinative morphology (e.g. the Turkish case paradigm), even a one-dimensional table suffices to represent the set of values that each morphological category can take. If two morphological categories are cumulated in a morpheme, as e.g. person and number in many conjugation paradigms, a two-dimensional table becomes necessary. If more than two categories are fused in a morpheme, or if inflection classes are involved in addition, a suitable combination of two-dimensional tables is chosen, as also shown in T8.

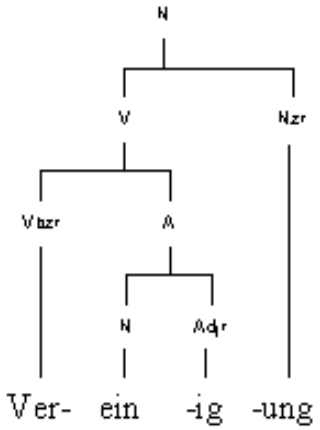
Although such complex paradigms are not necessary for agglutinative morphology, it is prudent at least to give a couple of examples of morphologically complex word forms to render the formulas concrete.

4.4 Syntagms

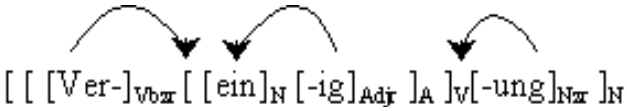
The accurate statement of the combination of a morphological marker (of a certain category) with its host may require specification of the category of the base, the conditions for the combination and specification of the resulting category. This may take the form of a rule or **instruction** how to construct a certain morphological form. E.g., concerning the Latin supine II like *dictū* 'to be said': in order to make the supine II of a verb, form its passive participle, convert this into a stem of the *u*-declension and form the ablative singular of this stem.

Syntagmatic relations are commonly analyzed in terms of constituency or dependency and accordingly visualized by **tree diagrams** or **dependency stemmata**, each illustrated here with an example:

S1. *Constituent structure of complex stem*



S2. *Dependency structure of complex stem*



The **valence class** of a verb or some other relational word is, above all, a syntactic property. However, in many languages, valence classes coincide with inflection classes, so that specification of a valence class may be relevant for morphology, too. This may be done in the form of a construction formula, e.g. for a trivalent verb: $[NP]_{Sbj}$ __ $[NP]_{dir.Obj}$ [AdvP]. Such valence classes may also contribute to an account of the function of a given case.

There are two basic types of **sequential order** of elements, scope order and template order. **Scope order** presupposes a binary construction consisting of a head and a dependent one of which has scope over the other, and specifies the position of the dependent as against the head. Thus, the position of an adnominal relative clause is determined with respect to its head nominal, and the position of the complement NP with respect to an adposition. At the morphological level, this convention is replaced by specifying the position of bound morphemes with respect to their hosts, or of grammatical formatives with respect to a lexical root or stem. In a syntagm such as S2, the relative order of the suffixes *-ig* and *-ung* is a contingent consequence of the fact that each of them has scope order with respect to the operand preceding it. This may be summarized in a formula like $[[X]_{N/V}$ *-ig* $]_A$.

Template order is the relative order of elements which do not form a construction with each other. It is typical of clitics which attach to whatever their host is, in a fixed order that does not reflect any semantic scope. Therefore template order cannot be specified in the same straightforward way as scope order. Instead, a schema is set up consisting of a fixed sequence of slots, a **slot** being a syntagmatic position for an element. For instance, the order of the prefixes on the Swahili verb may be described by the template in S3.

S3. *Swahili verb prefixes*

position	5	4	3	2	1	0
category	negation	subject	tense/aspect	relative	object	stem

The interpretation of such a template is straightforward: Each of the five prethematic slots is reserved for an element out of the paradigm designated by the slot label. If a slot remains unoccupied, this does not concern the order of the other elements.

Some morphological systems are so complex that different ranks of complex stems may be assumed such that affixes may be added at each rank. The suffix positions of the Kayardild noun may be schematized as in S4 (according to Evans 1995:122):

S4. *Kayardild nominal suffixes*

rank	0	1	2	3	4
category	stem	ADN/NUM*	relational$^{0-2}$	modal$^{0-2}$	(associating/complementizing)

At the first rank, the stem may be followed by any number of adnominal case and number suffixes. At rank 2, up to two relational cases, and at rank 3, up to 2 modal cases may follow. At rank 4, there may be either an associating or a complementizing suffix.

Templates are necessary when the sequential position of elements is not in consonance with their semantic scope. Structuralist accounts of the middle of the 20th century often give templates even for constructions which are decomposable into binary constructions in each of which scope order obtains. An example would be 'stem – past tense – personal ending' for the German regular finite verb. This is helpful in visualizing a complex form as a whole, but should be preceded by a separate account of each of the elementary constructions involved.

4.5 Illustrative examples

Abstract or formal statements should be supplemented by example expressions from the language being described. One of the most conspicuous differences between older and modern linguistic descriptions is the growing extent to which this principle is heeded. One might suppose this to be a trivial issue. Actually, however, descriptions that contain no examples are very often not intelligible and, to this extent, worthless. Since most linguistic statements fail in one or another respect and simply do not do justice to the real complexity of the language, an example can teach more than an ever so neatly formalized rule.

Above a certain degree of syntagmatic complexity, such an example should be presented in the canonical trilinear representation (cf. Art. 169), i.e. the text should be provided with an interlinear morphemic gloss and an idiomatic translation. While this was practically never done in traditional descriptions and even in structural linguistics, modern grammatical descriptions, especially those written in a functional or typological framework, often abound in analyzed illustrative examples. In this respect, there has been a real progress in descriptive linguistics.

There is a certain amount of trade-off between the corpus specimina as presented in the appendix and the examples in the running descriptive text. First of all, extensive illustration in the description can be reduced and be replaced by references to the corpus. Generally, such segments from the corpus are copied as illustrative examples in the running text which represent frequent patterns and do not involve any additional complications irrelevant to the issue at hand. Second, not all of the levels of representation figuring in the edition of the text corpus need be repeated in an example of the running text. For many purposes and especially for simple examples, some representation of the significans (orthographic or morphological, i.e. phonological with morpheme boundaries marked) coupled with a translation as a shorthand semantic representation will be sufficient. Thirdly, wherever uniformity of a set of examples is required, for instance in a conjugation paradigm, examples will not (entirely) be drawn from the corpus, but will rather complement it.

Normally, the illustration of a morphological phenomenon does not require a whole example sentence. Very often, a word form or, at most, a phrase will do. Where a sentence appears to be needed to adequately exemplify a grammatical phenomenon, this is a hint at the syntactic rather than morphological nature of the phenomenon. Also, while a corpus sentence may appear to be a more natural piece of language use, it may also distract, by its complexity, from the morphological phenomenon that is at stake. The situation is, of

course, different in syntax. There the author bears heightened responsibility if she simplifies corpus sentences for illustrative purposes.

References

CHERUBIM, D. 1973: „Grammatikographie.“ In: ALTHAUS, H. P. & HENNE, H. & WIEGAND, H.E. *Lexikon der germanistischen Linguistik*. Tübingen: Niemeyer. 768-778.

Cole, Peter 1982, *Imbabura Quechua.* Amsterdam: North-Holland (Lingua Descriptive Studies, 5).

Comrie, Bernard 1978. *Aspect: An introduction to study of verbal aspect and related problems*. Cambridge: Cambridge University Press.

COMRIE, B. & SMITH, N. 1977, “Lingua Descriptive Studies: Questionnaire.” *Lingua* 42:1-72.

Gabelentz, Georg von der 1881, *Chinesische Grammatik. mit Ausschluß des niederen Stils und der heutigen Umgangssprache.* Leipzig: Weigel.

Gabelentz, Georg von der 1891, *Die Sprachwissenschaft. Ihre Aufgaben, Methoden und bisherigen Ergebnisse.* Leipzig: Weigel Nachf. (2. überarbeitete Aufl.: Leipzig: Tauchnitz, 1901).

Givón, Talmy 1993, *English grammar. A function-based introduction. 2 vols.* Amsterdam & Philadelphia: J. Benjamins.

Haspelmath, Martin 1993, *A grammar of Lezgian.* Berlin: Mouton de Gruyter (Mouton Grammar Library, 9).

Haspelmath, Martin 2003, "The geometry of grammatical meaning: semantic maps and cross-linguistic comparison." Tomasello, Michael (ed.), *The new psychology of language. Cognitive and functional approaches to language structure.* Hillsdale, NJ: L. Erlbaum; 2:211-242.

Humboldt, Wilhelm von 1836, “Über die Verschiedenheit des menschlichen Sprachbaues und ihren Einfluß auf die geistige Entwicklung des Menschengeschlechtes.” Reimpr.: Humboldt, Wilhelm von 1963, *Schriften zur Sprachphilosophie.* [=*Werke in fünf Bänden*, hrsg. v. A. Flitner und K. Giel, Bd.III]. Darmstadt: Wissenschaftliche Buchgesellschaft; 368-756.

Jespersen, Otto 1937, *Analytic syntax.* Kopenhagen: E. Munksgaard.

LEHMANN, CH. 1980, „Aufbau einer Grammatik zwischen Sprachtypologie und Universalistik.“ In: BRETTSCHNEIDER, G. & LEHMANN, Ch. (eds.) *Wege zur Universalienforschung. Sprachwissenschaftliche Beiträge zum 60. Geburtstag von Hansjakob Seiler*. Tübingen: Narr (Tübinger Beiträge zur Linguistik 145), 29-37.

Lehmann, Christian 2002, *Possession in Yucatec Maya.* Erfurt: Seminar für Sprachwissenschaft der Universität (ASSidUE, 10). 2nd rev. ed.

Lehmann, Christian 2002, *Thoughts on grammaticalization.* Erfurt: Seminar für Sprachwissenschaft der Universität (ASSidUE, 9). 2nd rev. ed.

Lehmann, Christian 2004, "Data in linguistics". *Linguistic Review* 21/3-4:000-000.

Lehmann, Christian 2004, "Funktionale Grammatikographie." Premper, Waldfried (ed.), *Dimensionen und Kontinua. Beiträge zu Hansjakob Seilers Universalienforschung.* Bochum: N. Brockmeyer (Diversitas Linguarum, 4); 147-166.

MCGREGOR, William B. 1990, *A functional grammar of Gooniyandi*. Amsterdam & Philadelphia: J. Benjamins (SLCS, 22).

Mosel, Ulrike 2002, "Analytic and synthetic language description." Ezawa, Kennosuke & Kürschner, Wilfried & Rensch, Karl H. & Ringmacher, Manfred (eds.), *Linguistik jenseits des Strukturalismus. Akten des II. Ost-West-Kolloquiums, Berlin 1998.* Tübingen: G. Narr; 199-208.

Mosel, Ulrike 2003, Grammaticography - the art and craft of writing grammars. Kiel: Seminar für Sprachwissenschaft der Universität (www.linguistik.uni-kiel.de/mosel/PDF/Grgraphy.pdf).

McQuown, Norman A. 1990, *Gramática de la lengua totonaca. (Coatepec, Sierra Norte de Puebla)*. México, D.F.: Universidad Nacional Autónoma de México (Colección Lingüística Indígena, 4).

Noonan, Michael to app., "Grammar writing for a grammar-reading audience." Payne, Thomas & Weber, David (eds.), *Grammar writing.* Amsterdam & Philadelphia: J. Benjamins (Typological Studies in Language).

Plank, Frans (ed.). 1992. Paradigms. *The Economy of Inflection.* Berlin & New York: Mouton de Gruyter (*Empirical approaches to language typology, 9*).

Shopen, Timothy (ed.) 1985, *Language typology and syntactic description.* 3 vols. Cambridge etc.: Cambridge University Press.

Zaliznyak, Andrey. 1967. *Russkoe imennoe slovoizmenenie. [Russian nominal inflection]*. Moscow: Nauka.

Christian Lehmann
Universität Erfurt

Elena Maslova
Stanford University