

Grammaticography: The art and craft of writing grammars

Ulrike Mosel

1. Introduction

While lexicography is a well established branch of linguistics, represented in specialized journals and handbooks, grammaticography – the art and craft of writing grammars – is not.¹ Only the history of grammar writing, the typology of grammars or isolated problems feature as topics in language specific, especially applied linguistic research.² In contrast to grammatical analysis, grammar writing is not taught in linguistics courses or described in textbooks. It is apparently taken for granted that once one has analyzed the language and understood its phonology and morpho-syntax, one is also in a position to translate the analysis into a grammar. There are only a few works dealing with the question of what kind of information a reference grammar should contain and how this content could or should be presented.³ But to my knowledge there is none that focuses on how grammars are made from the first phase of collecting and analyzing data till the final product, the grammar, and discusses how the various components of this process interact. The present paper addresses a number of problems which the users of grammars easily detect, but the writers are often not aware of. Even where they are, their attempted solutions often involve less than satisfying compromises between what they envisage as the ideal grammar and what they can achieve given the limited resources of money, time and labour.

Various projects have started developing electronic grammars (e.g. Zaefferer 1998), but since to date no grammar of this format has been completed, they will not be considered here. We will also not analyse the role grammars play in language documentation or the difference between language description and language documentation (cf. Himmelmann 1998 on this point).

2. The typology of grammars

This paper focuses on the writing of grammars of previously undescribed, mostly endangered languages, which in some respects differ from other grammars. Like dictionaries, grammars can be classified on the basis of various criteria (compare Svendsén 1993: 17–23):

1. standard vs. dialect/ substandard grammars, i.e. the language or language variety described in the grammar can be either (written) standard language or a spoken regional or societal substandard;
2. monolingual vs. bilingual grammars, i.e. the language used for description can either be the native or a foreign language;
3. prescriptive vs. descriptive grammars;
4. active vs. passive grammars, i.e. grammars which take the perspective of the speaker and explain how words and phrases are used to express oneself in the language vs. grammars which from the hearer's perspective give an analysis of sentences and help the user to understand the grammatical categories and constructions.
5. grammars for language specialists like linguists, or teachers and non-specialists with no training in linguistics;
6. grammars for meta-linguistic purposes (linguistic theory building, linguistic typology) and grammars for language learning.

According to this typology, grammars of previously undescribed languages are usually bilingual descriptive grammars of spoken vernaculars rather than written standard languages. They are mainly written for academic specialists engaged in linguistic research and, consequently, designed as passive grammars. Such grammars can, however, become the basis for teaching materials and thus contribute to the development of literacy and standardization, for example in the case of Bauer's Maori grammar (1993) which was rewritten in a more accessible format (Bauer 1997).

3. Data collection, analysis and description

Probably every grammarian has had the experience that collecting and analyzing language data and writing a grammar are two different things: you think your analysis is perfect, you know how the language works, you might even speak it fluently, but when it comes to writing up the grammar

you are faced with unforeseen problems. How are you ever going to get all you know about the language into a single book? Unless you have already strictly followed a prefabricated questionnaire during fieldwork and confine your language description to writing up the answers, you are confronted with two questions: how you organize the language description in a linear sequence of chapters, and which area of grammar to choose as the starting point of writing. The answers to these questions are not the same. The sequential order of the chapters will follow more or less one of the major traditions. They will be either arranged in an ascending or a descending way, i.e. either from the smallest linguistic units to the larger ones or in the reverse order.

Since many linguistic phenomena are not related to one another in such hierarchies, but are interrelated in an intricate network, every grammar must make use of cross-references. Consequently, the grammarian should write the chapters requiring the least cross-referencing first. Which chapters these are depends on the structure of the language and the grammarian's theoretical approach. For instance, if you describe an inflectional language like German and you consider the various kinds of inflections as category establishing, you can write the chapter on word classes in connection with the morphology chapter. But if your classification of words is based on syntactic criteria, it is more practical to write this chapter after you have written the relevant parts of the syntax, irrespective of whether the word class chapter precedes or follows the syntax chapter.

Although collecting data, grammatical analysis, and writing up the chapters of a grammar are different tasks, they cannot be entirely separated, because once you start writing, you will discover gaps or inconsistencies so that you need to collect and analyze additional data. I often questioned my capacity as a fieldworker when I realized that my data were not sufficient. But now I think that the reason also lies in the very nature of writing, because – at least to some extent – the process of writing shapes and reshapes your thoughts which inevitably leads to changes in your analysis.

4. Factors determining the size, content and format of grammars

The size, content and format of a grammar is not only determined by the characteristics of the language, but also depends on the authors' knowledge of the language and their theoretical background, the prospective users, the intended purposes, and the time-frame set for its compilation.

4.1. The authors

Ideally grammars are written by a team of a native speaker and a person who fluently speaks the language as a foreign language, and where both have undergone a thorough linguistic training in linguistic theory and fieldwork methods. Working from different perspectives, native speakers and non-native speakers focus on different linguistic phenomena and hence complement each other (cf. Ameka, this volume). But such a dream team is rarely found in reality.

In fact, the linguist who plans to write a grammar of a previously undescribed language usually does not know the language, whereas the native speakers he or she works with do not have any linguistic training. As a consequence, both parties will have to engage in a social relationship of knowledge exchange: while the native speakers will teach the linguists their language, the linguists will explain what describing a language is all about. The speech community may also ask them to help with compiling a dictionary or producing educational materials and thus contribute to capacity building. Although funding agencies often do not explicitly support such involvement, many researchers do comply with such requests, because they feel that they owe the community something in return for their hospitality and cooperation (Newman and Ratliff 2001).

4.2. The users of grammars and what they use them for

Since antiquity, grammars have been written simply for the reason that human beings are curious about how their most important tool of communication works. The more practical purposes have been to

1. preserve the knowledge of older varieties of the language to guarantee that future generations will be able to understand sacred texts or otherwise greatly valued literature;
2. standardize the language;
3. teach a dominant or otherwise important language.

With the rise of historical linguistics in the nineteenth century and the development of linguistic typology and language universals research, scholars started investigating languages and writing grammars for purely scientific reasons. In recent times, more and more people, not only linguists, have

come to realize that in view of the imminent loss of much of the word's linguistic diversity, describing previously undescribed minority languages is an urgent task, not only for the sake of linguistics and related fields of research, but also for the communities themselves and their language maintenance efforts. The latter aspect, however, implies that grammarians should not only think about the design of grammars for linguists, but also develop strategies of how such grammars can be transformed into grammars for non-specialists. One of the problems to solve is, for instance, that the prospective users are not familiar with linguistic terminology, so the grammarian should keep scientific terminology to a minimum and explain every term he or she uses in simple words (cf. Bauer 1997; Mosel and So'o 1997).

4.3. Time frame

Existent efforts at making explicit how a grammar of a previously undescribed language should be organized, e.g. in the form of designing questionnaires or databases, tend to focus on comprehensive grammars as a resource for future linguistic investigations (Comrie and Smith 1977; Mosel 1987). Within these considerations, the time factor is not an issue. While Ph.D. grammars are submitted 3–5 years after fieldwork has begun, the publication of a comprehensive reference grammar of 500–800 pages may take between 5 to 10 years, or even longer depending on how many other commitments the researchers have and how long it takes to find a publisher. (cf. the prefaces of Dixon 1972, 1977; Evans 1995; Foley 1991; Mosel and Hovdhaugen 1992).

If grammaticography becomes a recognized branch of linguistics, the time factor should become an issue. Nowadays most grammars of previously undescribed languages are written by Ph.D. and Post-doc scholars who are under severe time pressure to simultaneously learn and analyze the language and cannot aim for a comprehensive grammar which would take more than 5 years to accomplish. The more restricted the time frame is the more important it is to set priorities. But what are the priorities? The answer to this question depends on several factors such as the structure of the language or whether the language is also well documented by texts and a dictionary. We'll discuss this topic below in the section on the content of grammars.

5. The borderline between dictionaries and grammars

Jespersen already observed: "Grammar and dictionary ... overlap and deal with the same facts", and in connection with English auxiliaries "there can be here no hard and fast line between grammar and dictionary" (Jespersen 1924: 32, 43). Any borderline drawn between lexicon and grammar is not a priori given by the language, but is a linguistic construct, so that it may be difficult to decide where to accommodate a particular linguistic phenomenon. Thus it might be difficult to decide whether certain words are more adequately treated as content words in the dictionary or as functional words in the grammar (cf. Schultze-Berndt, this volume). A similar problem is posed by productive speech formulas which are more or less grammaticalized means of expressing temporal and aspectual relationships, e.g. *be going to do something*, (French) *venir de faire quelque chose* 'have just done s.th.', *être en train de faire quelque chose* 'be doing something', (substandard German) *am Tun sein* 'be doing s.th.'. Are they to be described in the dictionary or in the grammar? And if they are treated in the grammar, do they belong in the chapter on tense/aspect or the one on infinitive constructions?

In his works on Kalam and English, Pawley (1993 and elsewhere) showed that the "parsimonious grammar-lexicon model of language" cannot adequately deal with productive speech formulas. Being neither totally productive nor totally fixed they "have no place in conventional grammars and dictionaries" (Pawley 1991: 433). The problem of accommodating function words and productive speech formulas in the dictionary or the grammar relates to the fact that both grammaticalization and lexicalization are continua.

A different kind of interface between grammar and lexicon is found in the interaction between aspect morphology and the semantics of the verbal lexicon, between number or noun class morphology and the semantics of nouns, or between derivational morphology and the lexicon. Here the semantic properties of lexemes determine to varying degrees which kind of grammatical or derivational morphology the lexemes can be combined with and what kind of meaning the resulting forms have.

To what extent the interrelationships between grammar and lexicon are accounted for in the grammar or in the lexicon depends on the nature of the language and the size of the grammar. Without doubt, the interaction between the lexicon and grammatical morphology will play a more important role in the grammar than derivational morphology will, and among deriva-

tional morphemes those which have a direct impact on constructions, as for instance valence changing morphology, will be given preference over others."

6. The format of grammars

Dictionaries usually consist of three parts: the main part which contains the dictionary entries, and what lexicographers call the front and the back matter (Svensén 1993: 230–235). While the front matter contains all information that is necessary to successfully use the dictionary (organization and scope of the dictionary, the resources used, the purpose, the intended user group, etc.), the back matter may consist of appendices giving encyclopedic or practical information, such as place names, weights and measures, quotations, proverbs or useful phrases for travelers. The structure of the main part of the dictionary is defined by the order of the headwords and the format of the dictionary entries, the former being called the macrostructure, the latter the microstructure (Svensén 1993: 223, 120).

Drawing parallels between the structures of dictionaries and grammars may help us to become aware of the needs of grammaticography, as both disciplines investigate how information on a language can be provided in an easily accessible format.

6.1. Front matter and back matter

In modern grammars the preface or the first chapter informs the reader about the language, its genetic relationships, its socio-linguistic background, the variety or varieties of language described in the grammar, the theoretical approach, the sources of the linguistic data, the methods used in text analysis and fieldwork, and finally, it gives a list of abbreviations. As it informs the reader about the content of the book rather than about the grammatical structure of the language, this introductory part clearly parallels the front matter of a dictionary.

Apart from the indices and the references, the last part of a grammar is often a short collection of texts. Corresponding to the back matter of dictionaries, the text collection does not fit into the organization of the content of the main part; it is an appendix which helps the reader to understand how the language is used in various contexts. But unless there exists some litera-

ture in the language or a text collection beside the grammar this 'back matter' should be more than an appendix of a few pages (cf. the section on texts).

6.2. Macrostructure: ascending and descending models

In lexicography the term macrostructure refers to the relative order of the dictionary entries, i.e. whether each entry constitutes a paragraph on its own or several entries are combined in one paragraph, and how the headwords are ordered: alphabetically, thematically, by frequency, etc. Correspondingly, we can speak of the macrostructure of a grammar when we talk about the order in which grammatical phenomena are described.

In general, the grammar can be organized in two directions as the units of grammatical analysis and description can be arranged in an ascending or descending order by either starting with the sound system and then moving to increasingly complex units (word, phrase, clause, sentence) or by starting from the sentence (or an even higher unit) and then moving down to smaller units. The ascending model (phonology > morphology > syntax), which probably has its origin in Priscian's Latin grammar (Hertz 1855–1859), seems to be preferred by most grammarians. Only the *Lingua Descriptive Series* grammars strictly follow the descending model: (Syntax (sentences > clauses > phrases) > Morphology (inflection > derivation) > Phonology). Other grammars, as for instance the *Samoan Reference Grammar* (Mosel and Hovdhaugen 1992), combine features of both the ascending and the descending model, as shown in Table 1.

Table 1. The structure of the Samoan Reference Grammar

Chapter	Content
1	Introduction ("front matter")
2	Phonology, orthography
3	The sentence, a preliminary view
4, 5	Word classes, Morphology
6–8	Types of phrases (noun phrase etc.)
9–14	Types of clauses (basic verbal clause etc.)
15–16	Sentences with embedded and dependent non-embedded clauses
17	Coordination (of words, phrases, and clauses)
18	Case marking and grammatical relations
	Indices, references ("back-matter")

Following the ascending model, we put the phonology right at the beginning because readers usually want to know how to pronounce and read words and sentences. But instead of presenting the morphology after the phonology, we jumped ahead to sentence structure to give the readers an idea of how the language works before they turn to any chapter of their particular interest. Since the book is a reference work, we do not expect the users to read the book from cover to cover, and therefore present each chapter as far as possible as a self-contained unit. The overview of sentence structure in the third chapter also helped us to avoid too many repetitions in the other chapters.

Then in the fourth and fifth chapter, we return to the ascending order and describe the word classes and their morphology, and subsequently the larger units of phrases, clauses and complex sentences. In the last two chapters we deal with phenomena which are not confined to a single level of description.

The three models presented in this section are far from ideal. All of them put phonology into one chapter although prosodic phonology pertains to all levels of grammar. Another problem area which was already mentioned above is that such models cannot cope with the inter-dependent relationships between the lexicon and the grammar of the language.

6.3. Microstructure: the presentation of form-meaning relationships

In lexicography, the microstructure is the format and internal design of the dictionary entry. A related issue in grammaticography would be the way examples are presented. It has become common practice to provide examples not in the running text, but on separate lines with several tiers, presenting the phonological or written form of examples and their meaning, with various numbers of intervening tiers for close phonetic transcription, morphemic glossing, and so on. Depending on the type of language and the purpose of the examples, the phonological or graphic form can be presented by various kinds of transcriptions, orthographies, or scripts and transliterations, whereas the meaning can be rendered by a free translation and in addition by an interlinear or literal translation.

The inventor of the technique of interlinear glossing is unknown; the earliest record I am aware of is the edition of Tolai texts by Meier (1909), which contains two stories with an interlinear word by word translation. In spite of its long tradition and its pervasive application, interlinear glossing has neither received attention in (meta-)grammaticography beyond Lehmann's article (1982) and the guidelines by editors, nor is it usually discussed in text books on grammatical analysis. The only exception I am aware of is Haspelmath (2002).

Interlinear glossing deserves more attention both in research and in teaching linguistics. Otherwise students might take it as an accurate representation of form-meaning relationships, which it is certainly not. The meaning of words and larger units of grammatical analysis does not equal the sum of the meanings of their component parts, as interlinear glossing might suggest, but results from the interaction of the meaning of the construction as such and the meanings of its parts. Thus interlinear glossing should only be seen as a tool to help the reader to understand examples; and useful tool indeed it is. This is not to say that interlinear glossing could not be further developed to become a means of representing form-meaning relationships, but at the moment grammaticography has taken only the first steps in this direction (Drude 2002).

Another future task of grammaticography could be to analyze to what extent graphs, figures and tables not only illustrate grammatical analysis, but also shape our view of language. How much did, for instance, the suggestive tree-diagram of immediate constituent analysis, which did not permit intersecting branches, influence the development of generative linguistics?

6.4. Terminology

Probably nothing has shaped our thinking about language more than the terminology we inherited from the Greeks and Romans. A case in point are the parts of speech, especially the distinction between nouns and verbs. For centuries these categories have been taken for universals; only a few linguists working on Amerindian and Austronesian languages have expressed their doubts (for a brief summary cf. Sasse 1993). But even those who are convinced that these categories do not exist in the language they describe have difficulties doing without the traditional terms, and due to the lack of an appropriate terminology, may express themselves in a contradictory way.

A second problem of traditional terminology is that it does not always clearly distinguish between form and function. A frequently found example is the confusion of word class and syntactic function when the term *adverb* is indiscriminately used for both and it is said that an adjective or a prepositional phrase may function as an adverb.

Thirdly, the same terms are used in various senses across linguistic traditions and theoretical approaches and, of course, also denote different things in different languages. Therefore, it is advisable to always define what grammatical phenomenon the terms denote in the grammar in question.

7. The content of grammars

Standard grammars of a previously undescribed language consist of more than just a description of the grammar of the language. Usually they also contain a chapter on the phonology in the beginning of the book and a collection of texts in the end. Some also include a vocabulary. In the following, we first deal with the content of the grammatical section and then discuss the role of texts in grammars.

7.1. Theory and practice

Theoretical discussions on the content of grammars aim at comprehensiveness (Comrie and Smith 1977; Lehmann 1989; Zaefferer 1998). Such considerations help us to make explicit what grammar writing

is all about, but in practice we must be realistic and decide on priorities, especially when grammars are written under the heavy constraints of Ph.D. or post-doc projects.

To some extent at least, the content of reference grammars reflects special interests of the authors and/or the editors and prevalent topics in current linguistic theories, such as subjecthood in the seventies or complex predicates in the nineties. The *Lingua Descriptive Series* very nicely mirrors the spirit of the times: it contains 79 questions on reflexives, but only five on negation. Although our linguistic knowledge is continuously expanding and linguistic fashions are changing, there are certain types of grammatical phenomena knowledge of which is essential for the understanding of a language. Knowing these phenomena forms the basis for the language learner as well as for the theoretician (theoreticians should also be language learners!):

1. the basics of the sound system and the orthography;
2. the structure of simple declarative, interrogative and imperative clauses;
3. the structure of word groups (NPs, verb complexes, etc.);
4. the formal features which indicate the syntactic functions of words and word groups (constituent order, case marking, cross-referencing etc.)
5. the classification of major word classes;
6. paradigms of inflecting words and the meaning of grammatical categories.

In principle, these are the things which the learner needs to know to understand simple texts with the help of a dictionary and which should be included in a sketch grammar. Once a grammarian has analyzed and described these basics, he/she can expand on them investigating more complex structures or the subtleties of simple constructions.

7.2. The role of texts in grammars

The low prestige of text editions (if they were more highly valued, linguists would no doubt publish more) can be attributed to several factors:

1. the politics of mainstream linguistics departments, some of which do not even recognize descriptive grammars as Ph.D. theses;

2. the fact that linguistic typology concentrates on the investigation of grammatical phenomena which manifest themselves in single sentences;
3. the fact that many typologists work with large samples of languages which does not allow the time consuming in-depth study of texts.

For scientific reasons, however, this relegation of texts to marginal appendices is not justifiable. Firstly, the grammarian like any other scientist should provide evidence for his or her claims, i.e. allow them to be independently verified. I suspect current linguistic practice in this regard is largely unparalleled in other sciences: Imagine a zoologist describing and analysing the parts of an animal without giving his colleagues the chance to see samples of it or at least photographs and films depicting the animal in full from various perspectives. Secondly, a text collection would give colleagues the opportunity to discover grammatical phenomena the linguist did not recognize, or did not have time to cover. In my comparison of Tolai and Tok Pisin (Mosel 1980), for instance, I overlooked the similarity between the Tolai particle *iat* and its Tok Pisin equivalent *yet*. I also did not describe the function of *iat* in my book on Tolai syntax (Mosel 1984), nor was it mentioned in Rinderknecht's Ph.D. thesis on the Tolai noun phrase (Rinderknecht 1987). But on the basis of my text edition (Mosel 1977), Sankoff (1993) was able to identify Tolai *iat* as a focus marker and relate it to Tok Pisin *yet*. This nicely illustrates how text editions can compensate for a grammarian's oversights. Thirdly, grammars are always written from the perspective of prevalent interests in contemporary linguistics. Only texts will later provide data for whatever had not been a topic of current linguists.

To conclude, every grammar should include or be accompanied by a collection of texts which enables the reader to understand how certain constructions are used in context and discover phenomena which were not described in the grammar. Ideally this text edition consists of annotated digitalized recordings of different language genres (e.g. myths, anecdotes, procedural texts, casual conversation, political debates and ritual speech events), accompanied by a transcription, a translation and a commentary on the content and linguistic phenomena.

8. Perspectives of description: Semasiological and onomasiological approaches

The form-meaning relationships of linguistic expressions can be analyzed and described from a semasiological or an onomasiological perspective; i.e. you can either investigate and describe the semantic properties of particular linguistic forms, or the various ways in which particular meanings are expressed. In spite of the works of Gabelentz (1891) and Jespersen (1924), the distinction between semasiological and onomasiological language description has not received much interest. Grammars are either semasiological or represent an arbitrary mix of these two approaches. Therefore, it seems worthwhile summarizing Gabelentz' and Jespersen's work in the following two sections (8.1 and 8.2). This will be followed by a brief discussion of the organization of the influential *Lingua Descriptive Series Questionnaire* (8.3). Finally, we argue that the semasiological and the onomasiological approach need to be separated (8.4) and summarize the problems of the onomasiological approach (8.5).

8.1. Gabelentz

One of the first scholars who thought about what linguistic competence means and how this competence could be described was Georg von der Gabelentz, who wrote in 1891 (Gabelentz 1984: 84):

I know a language firstly means: I understand it, when I hear or read it, and secondly, I use it correctly when talking or writing. Understanding the language means that it appears to me as a phenomenon, or better as a whole (system) of appearances, which I interpret. When using it, language is a means, or better a whole (system) of means of expressing my thoughts. In the former case the form is given and the content, the thought (the content expressed by the form) has to be found, whereas in the latter case the content of the thought is given and the form, i.e. the expression, has to be found.

Each linguistic expression can be viewed from two perspectives: the perspective of the hearer who analyses what he or she hears and the speaker who puts his or her thoughts into words. This idea then leads Gabelentz to the conclusion that a language consists of two interacting systems and that these systems should be described separately (Gabelentz 1984: 85): Since "everything in language is a phenomenon to be interpreted and a means to

be used"⁶, the ideal description of a language would present it as "two complementary grammatical systems"⁷:

the first one I call the analytical system, because it explains linguistic phenomena by breaking them down; the other one I call the synthetic system, because it shows how the grammatical means are made use of to construct speech.

Accordingly, an ideal grammar consists of two mutually complementary parts: the analytical or semasiological part takes the linguistic expressions as given, analyses their forms and describes their meanings, whereas the synthetic or onomasiological part takes the meaning as the point of departure and describes by what kind of linguistic forms they are expressed. The semasiological part precedes the onomasiological one, "because one needs to know how to interpret linguistic phenomena, before one is able to use the means of expression"⁹ (Gabelentz 1984: 86). These two parts are preceded by the description of the phonology.

Since the semasiological part takes the hearer's perspective, "the analysis has to start with the sentence, proceeding from the whole to the parts, i.e. from the sentence to the words and wordforms, eventually reaching the terminal elements, the single sounds"¹⁰ (Gabelentz 1984: 86). As for the organization of the onomasiological part, Gabelentz is less explicit. He only says that it should have sections for everything which is expressed by grammatical means (Gabelentz 1984:100) and refers to his *Chinese Grammar*. This, however, does not reflect what he said earlier about the onomasiological approach, but describes the syntax of Chinese.

8.2. Jespersen

In his book *The Philosophy of Grammar* (1924), Jespersen, who had probably been influenced by Gabelentz work¹¹, distinguishes "three stages of grammatical treatment of the same phenomena, of three points of view from which grammatical facts can be considered, which may briefly be described as (A) form, (B) function, (C) notion." (Jespersen 1924: 56) All form elements are to be treated in a section Jespersen calls "morphology"; they comprise sounds and 'sound combinations', "word elements, then words, and finally word combinations" (Jespersen 1924: 41) and even "word-order" (Jespersen 1924: 44). The functional categories (B), which

are described in the "syntax", are grammatical categories such as number, case, tense, mood, voice, person, gender.

In English, for instance, suffixes (e.g. *-ed* in *handed*), stem alternation (e.g. *drank*) or suppletive forms (e.g. *was*) are forms (A) which have the grammatical function (B) of 'preterit', which in turn expresses more than one notional category (C): 'past time', 'unreality' (*I wish I knew*), or even in certain contexts 'future time' (e.g. *It is time you went to bed*) (Jespersen 1924: 56). The notion of 'future time' or 'futurity', on the other hand, is not expressed by a 'real future tense' as in French, but "by means of phrases which do not signify mere futurity, but something else besides" such as 'volition' 'destiny', 'uncertainty' and 'obligation' (Jespersen 1924: 50), e.g. *I shall start tomorrow, he will start tomorrow, I am to start tomorrow* etc. (Jespersen 1924: 45f). In a similar way, Jespersen distinguishes between plural as functional or grammatical category and 'plurality' as the notional category of 'more than one'. Plurality can be expressed not only by plural, but also by singular forms, e.g. *horses, the upper and the lower shelf* (Jespersen 1924: 188 ff). Other notional categories discussed by Jespersen are 'person' (as expressed by pronouns, cross-referencing affixes and possessive constructions in forms of address like *your highness*), 'animate and inanimate' (Jespersen 1924: 234 ff.), "comparison" (Jespersen 1924: 244), "time" (Jespersen 1924: 254 ff.), "direct and indirect speech" (Jespersen 1924: 290 ff.), various types of "utterances" and "negation" (Jespersen 1924: 322 ff.). In general, only those categories figure as notional categories in onomasiological language descriptions which have some grammatical correlate. Jespersen says:

... it is important always to remember that these are to have a linguistic significance; we want to understand linguistic (grammatical) phenomena, and consequently it would not do to set to work as if language did not exist, classifying things or ideas without regard to their linguistic expression. (Jespersen 1924: 57)

8.3. Lingua Descriptive Series

The Lingua Descriptive Series Questionnaire (Comrie and Smith 1977) has provided the framework for the writing of thirty grammars of structurally and genetically very different languages. In accordance with the questionnaire, all the grammars have the same descending organization and use the same numbering of chapters, sections and subsections, so that cross-

linguistic comparison is made easy. Chapter 1 Syntax starts with (1.1) sentence types and subordination, then (1.2.) deals with the internal structure of clauses and phrases, followed by various topics related to the syntax of clauses, and their arguments and adjuncts: (1.3) coordination, (1.4) negation, (1.5) anaphora, (1.6) reflexives, (1.7) reciprocals, (1.8) comparison, (1.9) equatives, (1.10) possession, (1.11) emphasis, (1.12) topic, (1.13) heavy shift, (1.14) other movement processes. The chapter ends with (1.15) minor sentence types and (1.16) operational definitions of word classes. Chapter 2 Morphology comprises (2.1) inflection and (2.2) derivational morphology, and is followed by three short chapters on (3) phonology, (4) ideophones and interjections, and (5) the lexicon.

This radical descending organization is not user-friendly because it presents complex sentence structures before the reader is informed on the structure of simple clauses. Not surprisingly, Bauer changed this order when she rewrote her Maori LDS-Grammar (1993) into the more readable Reed Reference Grammar, now starting with an overview of simple clause and phrase structures including case marking, and describing complex sentence structures in the last part of the book (Bauer 1997: 5 ff., 540 ff.). Since the linguist's understanding of the structure of foreign languages begins with simple constructions (as does language learning in general), I am convinced that most people prefer grammars which start with simple constructions.

The questions of the questionnaire, and hence the grammars, mix the semasiological and the onomasiological approaches; "the structural framework of the Lingua Descriptive Studies is simply inconsistent, as it sometimes chooses a formal basis and another time a functional basis for its organization"¹², as Comrie (1998: 13) himself admits. Mixing both approaches can result in descriptions which fail to adequately account for the formal and semantic structure of languages. The section on structural questions, for instance, contains six questions concerning the existence of 1. verbs without subjects or dummy subjects, 2. verbs without direct objects, 3. a separate category of indirect object, 4. other kinds of arguments, 5. the combinations of different kinds of arguments, 6. the order of constituents. But it does not enquire about the argument structure of basic clauses in terms of coding properties.

In accordance with Jespersen's view of morphology, the means of expressing the 'syntactic functions' of 'subject' and 'object' are treated in chapter 2 'Morphology' in the section 'Noun inflection'. The beginning of this chapter is semasiologically oriented. One finds questions about the

various kinds of argument encoding devices including “pre-/postpositions” and “word order”, followed by questions on the encoding of agentive and non-agentive subjects with intransitive and transitive verbs and finally of subjects in copular constructions. The next question concerns the marking of objects and complements, but here without distinguishing semantic roles, the encoding of the agent in passive, pseudopassive and impersonal constructions, and where relevant the expression of topic and emphasized elements. Subsequently all these questions on syntactic functions are to be applied to “all types of non-finite or nominalized verb”.

The following section switches to onomasiological questions and asks the grammarian to indicate how non-local semantic roles are expressed. The roles of agent and patient are missing here. Obviously it is assumed that the expression of agents and patients is already exhaustively covered by the immediately preceding questions. But this is not so, and consequently other kinds of agent expressions cannot be accommodated without serious contradictions. In the Polynesian language of Tuvalu, for instance, “the agent of a transitive verb can be expressed as a modifier of a direct object” (Besnier 2000: 283).

- (1) *Ne kkati telotou niu*
 Pst cut their-3 coconut
 ‘They cut down the coconut tree’

For the lack of any other place, Besnier describes this construction in the section ‘Subject of transitive verb’, though the agent in this construction is an adnominal modifier and not a subject.

A grammar which keeps the semasiological and the onomasiological perspectives apart could deal with such phenomena in a non-contradictory way. In the semasiological part, it would describe the form and meaning of the argument structure and the construction of adjuncts including adnominal modifiers, whereas the various ways of expressing agents would be a topic in the chapter on the expression of semantic (thematic) roles in the onomasiological part.

8.4. Two grammars of the same language?

Gabelentz and Jespersen were convinced that languages should be described from both the semasiological and onomasiological perspective, but

actually did not provide any arguments for their position. So, we must ask whether both approaches are really necessary to capture a language. Reference grammars always choose the semasiological perspective – if not throughout, at least for the most part. This is justified and absolutely necessary because the readers need to know what kind of linguistic forms exist before they can understand that a particular meaning M1 is expressed by construction C1. Furthermore, grammatical constructions are often polysemous so that a particular construction C2 can have the different meanings M1, M2, M3. The onomasiological approach would describe how each of the meanings M1, M2, M3 is expressed in separate sections and possibly say that M1 is expressed by C1 and C2, whereas M2 is expressed by C2, C3, C4, and M3 by C2 and C5. This is shown in Figure 1.

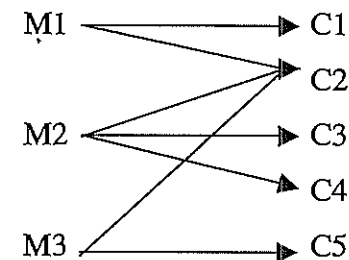


Figure 1. The onomasiological approach

If this onomasiological grammar has an index of construction types C1–C5, the readers can easily find the three meanings of C2. But what they cannot find out is how the different meanings are related to one another. Is there anything the meanings have in common? Is there a basic meaning from which the others are derived, perhaps by metaphor or metonymy?

A good example which shows how the various meanings of a polysemous construction are interrelated is found in Tolai. The most frequently used preposition in Tolai is *ta*. Prepositional phrases introduced by *ta* can refer to a place, a goal, a source, a point of time, a beneficiary, a recipient, an addressee, and a cause. An onomasiological description would describe these meanings in different sections and not reveal that the basic function of *ta* is to mark noun phrases (other than place names) as referring to places and that all other functions are derived from this locative function (Mosel 1984: 178–183, 191–193, 201–203; Lyons 1977: 718ff). What kind of entity or phenomenon a noun phrase marked by *ta* refers to, depends on the semantics of the context. For example, with verbs denoting motion towards a

goal, it is the goal, with verbs denoting motion away from something, it is the source, with verbs of giving, it is the recipient and with speech act verbs the addressee; e.g. *ruk ta* 'go into, enter', *vilau ta* 'run away from', *tar ta* 'give to', *biti ta* 'say to'.

Having described a language from the semasiological perspective and ideally written a comprehensive grammar which covers all grammatical phenomena and their meanings, the question arises: what is left for the onomasiological description? To date no descriptive reference grammar comprises the two parts of a semasiological and onomasiological description as envisaged by Gabelentz (1891), Jespersen (1924), Lehmann (1980, 1989) and myself (1987). But reference grammars occasionally contain single chapters or appendices with onomasiological descriptions. Most frequently such descriptions relate to the semantic function of possession, orientation in space and time, and negation, and the pragmatic function of questioning. The reason for this preference is obvious. These semantic and pragmatic functions play an important role in communication, but they cannot be treated together in single chapters of the semasiological description because they are expressed by linguistic constructions which pertain to more than one morpho-syntactic category or level. In addition, the onomasiological approach can account for productive speech formulas which partly fulfill grammatical functions.

The Samoan expression of temporal relations may serve as an example here. Samoan has six simple so-called TA-markers (tense/aspect particles) in the verb complex which relate the state of affairs talked about to the moment of the utterance or some other point of reference set by the context, e.g. *sā 'ai* 'ate (imperfective past)', *na 'ai* 'ate (perfective past)', *'ua 'ai* 'has eaten', *'olo'o 'ai* 'is eating', *e 'ai* 'will eat, usually eats', *'ole'ā 'ai* 'will definitely eat' (Mosel 2000). In addition to these TA-markers, Samoan has several temporal and aspectual productive speech formulas which complement the paradigm of TA-markers, as in Table 2.

While in the semasiological Samoan Reference Grammar these speech formulas are treated in the chapter on complement clauses, they would be put together with the TA-markers in a chapter on time in an onomasiological grammar. We may even go a step further and include other productive speech formulas which do not show any formal peculiarities like the complement taking aspectual and phasal verbs in Samoan, but are structurally similar to free constructions such as that exemplified in Table 3.¹⁴

Table 2. Samoan temporal and aspectual productive speech formulas

TA	TEMP/ASP.V	COMP	EVENT	gloss
	<i>leva</i>	<i>ona</i>	_____	'be a long time ago that _____'
	<i>lata</i>	<i>ona</i>	_____	'be close that _____, soon'
	<i>'uma'</i>	<i>ona</i>	_____	'be finished that _____, already'
	<i>'āmata</i>	<i>ona</i>	_____	'start to _____'

Table 3. Other productive speech formulas

<i>na alu</i>	_____	<i>'ātoa</i>	PROG	_____			
TIME SPAN			EVENT				
<i>Na</i>	<i>alu</i>	<i>le</i>	<i>aso</i>	<i>'ātoa</i>	<i>'o</i>	<i>o'u</i>	<i>moe</i>
PAST	go	ART	day	whole	PROG	1.SG	sleep

'I slept the whole day.' (lit. 'The whole day I was sleeping went.')

Do these speech formulas belong in a grammar or a dictionary? There is no straightforward answer. For the grammarian the answer largely depends on the purpose of the grammar. If the grammar is meant to be a book about constructions people use to express certain types of meaning, then these expressions belong there. But if the grammar is only concerned with morphology and syntactic rules of phrase and sentence structure, they will be ignored because they do not formally differ from non-formulaic expressions. Therefore such formulas were not included in the Samoan Reference Grammar, but in the language course book *Say it in Samoan* (Mosel and So'o 1997).

The only onomasiological grammar I have come across is *A communicative grammar of English* by Leech and Svartvik (1975), which complements the English reference grammar by Quirk et al. (1972). In this grammar one finds, for instance, a chapter on "time, tense and aspect" (Leech and Svartvik 1975: 63 ff), with a section on 'future time', which starts stat-

ing that “there are five ways of expressing future time in the English verb phrase. The most important future constructions are those which use *will* (*shall*) and *be going to*.” (Leech and Svartvik 1975: 70)

8.5. Structure and content of onomasiological descriptions

A tradition of 2000 years has shaped the structure of semasiological grammars and made linguists more or less intuitively agree on their contents. Onomasiological descriptions, on the other hand, have yet to find an appropriate, widely accepted design. There are too many questions which have not yet been answered:

1. How are the semasiological and the onomasiological description to be related to each other? Should the onomasiological description just be an addition to the semasiological one or should it be more independent?
2. How can we decide on which semantic and pragmatic domains are to be chosen for the onomasiological part? Would these domains all have the same status? How are the descriptions of the domains to be ordered in the grammar? Would they be the same for all languages?
3. To what extent will the descriptions account for lexical and phraseological aspects of the linguistic phenomena to be described?

Linguistic typology and cognitive linguistics will certainly bring us nearer to an answer, but the task of transferring their findings into the practice of writing grammars should not be underestimated. As with traditional grammar writing, it requires setting priorities and organizing the description of very complex phenomena into a sequence of chapters in a book.

9. Concluding remarks

In view of the urgent necessity for documenting endangered languages, grammaticography needs to be advanced as a genuine discipline of linguistics to ensure that the analysis of previously undescribed linguistic phenomena is presented in such a way that it can serve as the basis for further research as well as the development of materials for language maintenance measures. In the preceding chapters I first compared grammaticography with its older sister lexicography. In the search for methodologies of ‘catch-

ing language’, grammarians can certainly learn from lexicographers in how to present information on a language in a format that fulfills the standards of the trade, but is user-friendly at the same time.

The second part briefly discusses the problem of how to decide on the content of a grammar and the importance of texts as complementing the grammatical description. Instead of working with a workplan which aims at a comprehensive grammar, I suggest starting with a short grammar comprising the essentials and then, if time allows, expand on topics of special interest.

The final part of the article first draws attention to the pioneers of grammaticography, Gabelentz and Jespersen, who distinguished between the semasiological and the onomasiological approach to language description. While the semasiological description analyses the forms of expression and their meanings, the onomasiological description shows the interaction of various kinds of lexical and grammatical means of expression to cover particular semantic and pragmatic domains like possession, orientation in space and time, or asking questions. Both approaches are necessary.

Writing a grammar is more than an organizational exercise. As Bob Dixon demonstrates in his grammars and in his teaching of future grammarians, it requires both a deep understanding of language in all its complexity and a sound theoretical basis (cf. Dryer, this volume). In the conclusion of his book on ergativity, Dixon (1994: 229) remarks,

The most important point is that a language can only profitably be studied as whole. One must recognize and distinguish different levels of structural organization – phonological, morphological, syntactic, semantic, discourse and pragmatic – but each of these continuously interrelates with the others.

Here lies the challenge and the fascination of writing grammars, and the task of grammaticography as a future discipline of linguistics. In a world where specialists – and linguists are no exception – know more and more about less and less, it becomes increasingly important to develop methodologies of making specialized knowledge accessible to non-specialists. Only the identification, analysis and description of the essentials of the structure of languages will enable us to connect specialized knowledge of various linguistic areas and advance our understanding of language. For this very reason the old tradition of grammar writing is gaining more importance than ever.

Notes

1. I am grateful to Philipp Brandenburg, Alan Dench, Nick Evans, Geoffrey Haig and Nicole Nau for their valuable comments on earlier versions of this paper and inspiring discussions on grammaticography.
2. Ágel (1995 ed.); Cherubim (1973); Helbig (1992); Leitner (1986); Wilkins (1976).
3. Cf. Comrie and Smith (1977); von der Gabelentz (1901); Jespersen (1924); Lehmann (1980); Mosel (1987); Nida (1949: 222–281).
4. Cf. Vonen's criticism of the Samoan Reference Grammar which is quite justified with regard to the inconsistency of terminology (Vonen 1997: 139–144, Mosel and Hovdhaugen 1992).
5. "Ich kann eine Sprache, das heißt erstens: ich verstehe sie, wenn ich sie höre oder lese,- und zweitens: ich wende sie richtig an, wenn ich in ihr rede und schreibe. Insofern ich sie verstehe, stellt sie sich mir dar als Erscheinung, oder richtiger, als eine Gesamtheit von Erscheinungen, die ich deute. Sofern ich sie anwende, bietet sie sich mir als Mittel, oder richtiger als eine Gesamtheit von Mitteln zum Ausdruck meiner Gedanken. Dort war die Form gegeben und der Inhalt, der Gedanke zu suchen; hier umgekehrt: gegeben ist der Gedankeninhalt, und gesucht wird die Form, der Ausdruck."
6. "Alles in der Sprache ist zugleich zu deutende Erscheinung und anzuwendendes Mittel".
7. "zwei einander notwendig ergänzende grammatische Systeme"
8. "das eine nenne ich das analytische, weil in ihm die Spracherscheinungen durch Zerlegung erklärt werden; das andere nenne ich das synthetische, weil es lehrt die grammatischen Mittel zum Aufbau der Rede zu verwerthen"
9. "denn man muss die Spracherscheinungen deuten können, ehe man die Sprachmittel anwenden kann."
10. "...und so hat die Analyse vom Satze auszugehen. Folgerichtig schreitet sie vom Ganzen zu den Theilen, also vom Satze zu den Wörtern und Wortformen fort, und erst zuletzt gelangt sie zu den letzten Elementen, den einzelnen Lauten."
11. Cf. the footnote in Jespersen (1924: 39); Kürschner (2002).
12. My English translation from the German "Der strukturelle Rahmen der Lingua Descriptive Studies ist schlichtweg inkonsistent, indem er einmal eine formale und ein andernmal eine funktionale Basis für die Organisation wählt." This in turn is a translation from Comrie's English original by Christian Strömsdörfer and Dietmar Zaefferer.
13. Similar constructions are found in Samoan. (Duranti 1981: 173 f; Mosel and Hovdhaugen 1982: 761–763).
14. For further examples cf. Mosel (2002).

References

- Ágel, Vilmos and Rita Brdar-Szabó (eds.)
 1995 *Grammatik und Deutsche Grammatiken*. Tübingen: Max Niemeyer Verlag.
- Ameika, Felix
 2006 Real descriptions: Reflections on native speaker and non-native speaker descriptions of a language, [this volume.]
- Bauer, Winifred (with William Parker and Te Kareongawai Evans)
 1993 *Maori*. Descriptive Grammars Series. London and New York: Routledge.
 1997 *The Reed Reference Grammar of Maori*. Auckland: Reed Publishing Ltd.
- Besnier, Nico
 2000 *Tuvaluan. A Polynesian Language of the Central Pacific*. Descriptive Grammars Series. London and New York: Routledge.
- Cherubim, Dieter
 1973. Grammatikographie. In *Lexikon der Germanistischen Linguistik*, H.P. Althaus, H. Henne and H.E. Wiegand (eds.), 768–778. Tübingen: Max Niemeyer Verlag.
- Comrie, Bernhard
 1998 Strukturrahmen für Deskriptive Grammatiken. In *Deskriptive Grammatik und Allgemeiner Sprachvergleich*, Zaefferer, Dietmar (ed.), 7–16. Tübingen: Max Niemeyer Verlag.
- Comrie, Bernhard and Norval Smith
 1977 *Lingua Descriptive Studies: Questionnaire*, *Lingua* 42: 1–71.
- Dixon, R. M. W.
 1977 *A Grammar of Yidjñ*. Cambridge: Cambridge University Press.
 1994 *Ergativity*. Cambridge: Cambridge University Press.
 1997 *The Rise and Fall of Languages*. Cambridge: Cambridge University Press.
- Drude, Sebastian
 2002 Advanced glossing – A language documentation format and its implementation with shoebox. In *Proceedings of the International LREC Workshop on Resources and Tools in Field Linguistics (Las Palmas 26-27 May 2002)*, Peter Austin, Helen Dry and Peter Wittenburg (eds.).
- Dryer, Matthew
 2006 Descriptive theories, explanatory theories, and Basic Linguistic Theory, [this volume.]

- Duranti, Alessandro
1981 *The Samoan Fono: A Sociolinguistic Study (Pacific Linguistics B-80)*. Canberra: The Australian National University.
- Evans, Nicholas D.
1995 *A Grammar of Kayardild*. Berlin and New York: Mouton de Gruyter.
- Foley, William A.
1991 *The Yimas Language of New Guinea*. Stanford: Stanford University Press.
- Gabelentz, Georg von der
1984 Reprint. *Die Sprachwissenschaft. Ihre Aufgaben, Methoden und bisherigen Ergebnisse*. [3rd ed.] Tübingen: Gunter Narr Verlag. Original edition, Leipzig 1891.
- Helbig, Gerhard
1992 Grammatiken und Ihre Benutzer. In *Offene Fragen - Offene Antworten in der Sprachgermanistik*, Vilmos Ágel and Regina Hessky (eds.), 135-150. Tübingen: Max Niemeyer Verlag.
- Hertz, Martin (ed.)
1855-59 *Prisciani Grammatici Caesariensis Institutorum Grammaticarum Libri xviii*. Reprinted 1981. Hildesheim: Olms.
- Himmelmann, Nikolaus
1998 Documentary and descriptive linguistics. *Linguistics* 36:161-195.
- Jespersen, Otto
1924 *The Philosophy of Grammar*. Reprinted 1965. New York: Norton.
- Kürschner, Wilfried
2002 Otto Jespersen (1960-1943) und Georg von der Gabelentz (1940-1893). In *Linguistik jenseits des Strukturalismus (Akten des II. Ost-West Kolloquiums)*, Kennosuke Ezawa and Wilfried Kürschner, Karl H. Rensch and Manfred Ringmacher (eds.), 173-185. Tübingen: Gunter Narr Verlag.
- Leech, Geoffrey and Jan Svartvik
1975 *A Communicative Grammar of English (Pb. edition)*. London: Longman.
- Lehmann, Christian
1980 Aufbau einer Grammatik Zwischen Sprachtypologie und Universalistik. In *Wege zur Universalienforschung*, Gunter Brettschneider and Christian Lehmann (eds.), 29-37. Tübingen: Narr.
- 1982 Directions for interlinear morphemic translations. *Folia Linguistica* 16: 199-224.
- 1989 Language description and general comparative grammar. In *Reference Grammars and Modern Linguistic Theory*, Gottfried Graustein

- and Gerhard Leitner (eds.), 131-162. Tübingen: Max Niemeyer Verlag.
- Leitner, Gerhard (ed.)
1986 *The English Reference Grammar*. Language and linguistics, writers and readers. Tübingen: Niemeyer.
- Lyons, John
1977 *Semantics*. Cambridge: Cambridge University Press.
- Meier, Joseph
1909 *Mythen und Erzählungen der Küstenbewohner der Gazelle-Halbinsel (Neu-Pommern)*. Münster: Aschendorffschen.
- Mosel, La'i Ulrike and Ainslie So'o
1997 *Say it in Samoan (Pacific Linguistics D-88)*. Canberra: The Australian National University.
- Mosel, Ulrike
1977 *Tolai Texts. Kivung 10/1-2*. Port Moresby: Linguistic Society of Papua New Guinea.
- 1984 *Tolai Syntax and its Historical Development*. Pacific Linguistics B-92. Canberra: The Australian National University.
- 1987 *Inhalt und Aufbau deskriptiver Grammatiken*. Köln: Seminar für Sprachwissenschaft.
- 2000 Aspect in Samoan. In *Probleme der Interaktion vom Lexik und Aspekt (ILA)*, Walter Breu (ed.), 92-97. Tübingen: Niemeyer.
- 2002 Analytic and synthetic language description. In *Linguistik jenseits des Strukturalismus (Akten des II Ost-West Kolloquiums)*, Kennosuke Ezawa, Wilfried Kürschner, Karl H. Rensch and Manfred Ringmacher, 199-208. Tübingen: Gunter Narr Verlag.
- Mosel, Ulrike and Even Hovdhaugen
1992 *Samoan Reference Grammar*. Oslo: Scandinavian University Press.
- Newman, Paul and Martha Ratliff
2001 Introduction. In *Linguistic Fieldwork*, Paul Newman and Martha Ratliff (eds.), 1-14. Cambridge: Cambridge University Press.
- Nida, Eugene A.
1949 *Morphology: The Descriptive Analysis of Words*. 2nd ed. Ann Arbor: The University of Michigan Press.
- Pawley, Andrew
1991 Saying things in Kalam: reflections on language and translation In *Man and a Half (Essays in Pacific Anthropology and Ethnobiology in Honour of Ralph Bulmer)*, Andrew Pawley (ed.), 432-444. Auckland: The Polynesian Society.
- 1993 A language which defies description by ordinary means. In *The Role of Theory in Language Description*, William A. Foley (ed.), 87-129. Berlin/New York: Mouton de Gruyter.

- Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech and Jan Svartvik
1972 *A Grammar of Contemporary English*. London: Longman.
- Rinderknecht, Peter
1987 *Nomen und Verb im melanesischen Tuna (Tolai)*. Bern etc.: Peter Lang.
- Sankoff, Gillian
1993 Focus in Tok Pisin. In *Focus and Grammatical Relations in Creole Languages*, Francis Byrne and Donald Winford (eds.), 117–140. Amsterdam, Philadelphia: John Benjamins Publishing Company.
- Sasse, Hans-Jürgen
1993 Syntactic categories and subcategories. In *Syntax* (Handbücher der Sprach- und Kommunikationswissenschaft Bd. 9.1), Joachim Jacobs, Arnim von Stechow, Wolfgang Sternefeld and Theo Vennemann (eds.), 646–653. Berlin/New York: Mouton de Gruyter.
- Schultze-Berndt, Eva
2006 Taking a closer look at function verbs: Lexicon, grammar, or both?, [this volume.]
- Svensén, Bo
1993 *Practical Lexicography. Principles and Methods of Dictionary-making*. Oxford/New York: Oxford University Press.
- Vonen, Arnfinn Muruvik
1997 *Parts of Speech and Linguistic Typology*. Open classes and conversion in Russian and Tokelau. Oslo: Scandinavian University Press.
- Wilkins, D.A.
1976 *Notional Syllabuses*. London: Oxford University Press.
- Zaefferer, Dietmar (ed.)
1998 *Deskriptive Grammatik und allgemeiner Sprachvergleich*. Tübingen: Niemeyer.

Real descriptions: Reflections on native speaker and non-native speaker descriptions of a language¹

Felix K. Ameka

There are 2,000 or 3,000 languages, for which we have no decent description, that will pass into disuse within the next few generations. Trained linguists are urgently needed to document them. In most cases native speakers can be trained as linguists but in many instances an outsider is required. (Dixon 1997a:138)

1. Introduction

Now more than ever in the history of linguistics with our growing awareness of the need to document and describe languages that have not yet been recorded, the importance and role of the native speaker has been brought into sharper focus. The native speaker is no longer merely seen in field linguistics as “a human being who is able to give information about his or her language” (Coulmas 1981: 5). They are rather seen as people who can help “shape the record” (Mithun 2001) or promote the work of the linguist through a healthy and symbiotic relationship of the kind described by McLaughlin and Sall (2001). The native speaker may also be the last repository of a significant body of knowledge which is a crucial component of the world’s intellectual and cultural heritage (cf. Crystal 2000, Evans 2001). At the same time, it has become increasingly clear that there are different kinds of native speakers (see e.g. Dorian 1981, 1986 and Grinevald 2003; and see also Davies 2003 on the role of the native speaker in applied linguistics). Despite the recognition of a more active role for the native speaker in language description, discourses on documentation generally take it for granted that the linguist doing the documentation is an outsider—a non-native speaker of the language in question. Some will concede, as Dixon does in the quote above, that native speakers can also be trained

Trends in Linguistics

Studies and Monographs 167

Editors

Walter Bisang

(main editor for this volume)

Hans Henrich Hock

Werner Winter

Mouton de Gruyter
Berlin · New York

Catching Language

The Standing Challenge of Grammar Writing

edited by

Felix K. Ameka

Alan Dench

Nicholas Evans

2006

Mouton de Gruyter
Berlin · New York