Toward a balanced grammatical description

by Thomas E. Payne

1. Introduction

Scholars who attempt to write linguistic grammars of underdocumented languages strive to accomplish several worthy goals, many of which seem to conflict with one another. Tension and discouragement may arise as the author of a grammar attempts to balance often competing goals and values. Among the many tensions faced by grammar writers, the following stand out as being particularly perplexing:

- Inclusiveness vs. usefulness (publishability).

- Technical accuracy vs. understandability.

- Universality vs. specificity.

- Responsibility to the academic community vs. responsibility to speakers.

- A "form-driven" outline vs. a "function-driven" outline.

The descriptive linguist must balance all of these tensions (and more!) in an esthetically pleasing, intellectually stimulating, and genuinely informative package. Writing a linguistic grammar is definitely an art as well as a science. The present paper is largely an attempt to elaborate on and help field linguists resolve these potential tensions in their own contexts.

2. What is a grammatical description anyway?

Before looking at the specific tensions that writers of grammatical descriptions face, I would like to spend a few paragraphs considering the question of what it is we are creating when we write a linguistic grammar. I believe it is important to always keep the larger view in mind as we attempt any large and complex task, such as writing a grammar.

2.1. A grammatical description is a communicative act

Sometimes grammar writers tend to forget that a written grammar is an act of communication (see T. Payne 2007 for elaboration of this idea). The writer has important knowledge to share with a particular, interested audience. The description will succeed or fail to the extent that it communicates that knowledge in a way that the intended audience is able to appreciate and incorporate into their own cognitive frameworks.

Like any communicative act, a grammatical description has several characteristics:

- All communicative acts have a "speaker". The speaker in the case of a linguistic grammar is the descriptive linguist. Each linguist has particular

interests, goals, personality and background, all of which contribute to the unique characteristics of the grammar. For example, some linguists are very analytic and detail oriented, while others are more "global" in their approach to life in general, and grammatical description in particular. Detail-oriented individuals may relish the nuances of phonological and morphophonemic variation, but find syntax and pragmatics overwhelming to deal with at a level of detail that they are comfortable with. More global thinkers, on the other hand, may delight in making grand claims about discourse, syntax and "information flow," but have little patience with allophonic variation. Grammars written by different writers will reflect these kinds of individual predilections. There is no one correct grammar of a language, just as there is no one correct way to accomplish any speech act, such as apologizing, sermonizing, encouraging or proposing marriage.

- A grammatical description is simultaneously underspecified and redundant. The terms "impoverished" and "exuberant" respectively were used by Alton "Pete" Becker (see, e.g., Becker 1979) to describe these characteristics of communicative acts in general. The text itself cannot explicitly "encode" all the information necessary for it to be understandable. Much important information is necessarily left implicit, to be inferred by the audience (see also Grice 1975, Sperber & Wilson, 1995 and the other literature on the ostensive and inferential nature of human communication). At the same time, particularly salient information must be highlighted in special ways, and referred to multiple times throughout the text to keep it from simply blending into the background. This "texture" of highlighting new, important or asserted information, and downplaying old, background, and presupposed information is a feature of all successful communicative acts, and should be a feature of linguistic grammars as well.

- A grammatical description takes place in a context. Pragmatics is the study of linguistic context. Any communicative act makes assumptions about the audience, and attempts to engage them "where they are." Many of the important concepts in pragmatics refer to assumptions that people make about their interlocutors when engaging in communication, e.g., what the audience already knows ("given" information), what they are thinking about at the moment ("activated" information) and what is important to them ("newsworthy" information). The audience for a linguistic grammar is a community of linguists (often a dissertation committee) who are intensely interested in the details of individual languages, but who may not know much about the particular language being described. For this reason, background assumptions about general linguistic concepts can be assumed or treated lightly -- the text of an academic grammar need not provide a basic course in linguistics. However, it must explicitly and carefully highlight the data and perspectives that the author wishes to add to the general body of knowledge. Of course there are many kinds of grammatical descriptions, including school grammars, pedagogical grammars and others. The focus of the present paper is on linguistic, or reference grammars. The context for this type of grammar is usually a particular scholarly tradition that the grammar writer is a part of, often in a graduate degree program. Other types of grammars will incorporate

different assumptions about the backgrounds, interests, and states of mind of their readers.

All too often grammar writers tend to forget that a grammar is a communicative act. Linguists, of all people, should be aware of the properties of communicative acts in general, and should be able to apply this awareness to their own work whenever appropriate. For some reason, however, grammar writers often view their work as a schematic diagram of the categories and patterns that constitute the internal (subconscious) grammar of a language, or as a logical machine that "generates" or "sanctions" grammatical structures. While schematic diagrams and logical machines can be useful for certain purposes, they are very different kinds of objects than most human discourse. Therefore they tend to be unsatisfactory as communicative acts.

2.2. A grammatical description is a work of non-fiction literature

Perhaps I am old fashioned, but I still believe that the best model for descriptive grammars is a book that will eventually be available for use in libraries, and on night tables of many interested individuals. I realize, of course, that there are now many other possible models, primarily due to the explosion of computational technologies in the past few decades. History may eventually prove me wrong, nevertheless, I still recommend that grammar writers use the "non-fiction literature" model for their grammatical description, rather than any of the current more computationally oriented models, such as annotated databases, hypertext documents or expert systems. Even a grammar that is published on the internet, I believe, should have characteristics known to hold of good non-fiction literature. In particular, it should be a coherent whole with a beginning, middle and end (rather than just a system of cross-references to fragmented topics, as so many technical manuals or "help systems" tend to be). It should also be interesting and engaging to the intended audience. Toward this end, I would like to highlight the following desirable characteristics:

- A grammatical description should be generously illustrated with examples. Even if no one remembers or accepts the technical analyses in a grammar, future readers should still be able to use and interpret the data. Linguistic data (not theoretical points or frameworks) constitute the subject matter of a linguistic grammar. Consequently data should always be in primary focus. Data may be presented in charts, e.g., paradigms of verb forms or pronoun systems, examples elicited in imagined contexts, examples extracted from actual contexts, or entire transcribed and analyzed texts. A good linguistic grammar will employ all of these methods of presenting data, as they all have their particular functions in providing a full and satisfying portrait of a language.

- A grammatical description should start with simple, clear topics, and gradually introduce complexity and irregularity. I find there is a tendency for some grammar writers to state a pattern, and then immediately give the exceptions to the pattern, sometimes even before any straightforward examples! Perhaps this is because grammar writers are so aware of the complexities that the regular, simple and straightforward facts seem like "lies." It feels wrong to state a generalization knowing all along that there are many exceptions,

contradictions and variations. However, I must implore grammar writers to put themselves in the shoes of their audience for a moment. Exceptions and variation do need to be presented in due course. However, if there is a pattern worth stating, it is worth giving the audience time and opportunity to digest it, and to incorporate it into their own emerging "image" of the language before they are exposed to a complex range of exceptions. The same principle applies to the order in which topics are presented. Most good linguistic grammars start with "lower level", relatively regular topics -- phonology, morphology, noun phrase structure -- and work their way "up" to more complex topics, such as verb phrase structure, clause structure, clause combining, etc. This, in general, is a reasonable approach, though see Section #.# below for a caveat regarding a strictly structural organization of a grammar.

- Most of the exposition in a grammatical description should be in clear prose, rather than complex diagrams, charts and formulas. As mentioned above, formulas and diagrams can be helpful to a grammar writer in clarifying thoughts, and presenting knowledge in a precise way. However, over-reliance on formulas and diagrams can obscure rather than elucidate knowledge. This is especially true when the formulas and diagrams stem from a particular theoretical tradition. Theoretical traditions in linguistics are notoriously short-lived. What is currently in vogue will tomorrow be anachronistic. Much good linguistic work of past decades remains largely inaccessible to modern scholars simply because the frameworks employed have gone the way of the dinosaurs.

In summary, an important part of writing a "balanced" grammatical description is keeping the big picture in mind. Only if we remember what it is we are trying to do that we can do it in a way that is communicative and enjoyable to read. Grammar writers will do well to remember that the grammar they are producing is an act of communication and that it is a work of non-fiction literature.

3. Inclusiveness vs. Usability

All field linguists want to write an inclusive grammatical description. Field linguists typically have vast knowledge of the language they are describing, and it pains them not to express ALL of that hard-earned knowledge between the covers of the written grammar. However, a completely inclusive grammatical description has never been written, and if it were it could never be published. Valuable knowledge buried in reams of explicit but relatively minor detail becomes hard to recognize, and virtually useless. Even a thousand page tome cannot represent all the categories and habitualized patterns that make up the grammar of everyday talk. Therefore, all grammars must compromise inclusiveness to a certain extent in order to ensure that the grammar will be usable, and hence attractive to potential publishers.

I have known many fieldworkers who have felt paralyzed when it comes to describing the grammar of a language because they don't know where to start. They can't say *anything* until they can say *everything*. Unfortunately, the point will never come when you can say everything, so you might as well say what you know.

It has been observed by many researchers (e.g., Grice, 1981, Sperber and Wilson 1995 *inter alia*) that utterances are only partial representations of speaker intentions. Much of what is communicated via language is understood via inference. This is the normal way that humans go about the business of communication, and applies as well to a grammatical description, as I hope to show below. Sometimes saying too much can actually detract from the communicativity of a speech act. Here's an example from an actual conversation:

(1)   He's holding her hand the whole time across the table.

A lot of detail has been left out of this utterance. For example, most people have two hands, yet the sentence does not mention WHICH of "her hands" the subject is holding. In fact, if the speaker did specify "her right hand" it may be a potential distraction. The hearer may legitimately wonder why the speaker is mentioning her *right* hand. There must be some relevance to that detail, and the hearer, being a cooperative conversationalist, would try to identify the relevance of the right, vs. left hand.

Something similar is true in grammar writing. Saying too much may not only bury relevant information, but may actually confuse readers. For one small example, consider the issue of word classes, or "parts of speech". The importance of word classes has seldom been questioned in discussions of what should be included in a grammatical description of a language. How can you even begin to describe a language if you don't have a clear idea of what the building blocks of that language are? Every grammatical description must at least make mention of nouns and verbs, and probably adjectives, adverbs and some kinds of particles as well. Yet, with a little reflection, it becomes clear that classes such as "Noun" and "Verb" are no more than convenient approximations, rather than absolute categories. They are imprecise generalizations that help readers understand something important about a language, but which do not directly correspond to fixed categories in even one language.

If you investigate the grammatical properties of a number of words, you soon find that the lexicon of any language is not divided into clear, mutually exclusive classes. There are in fact very good examples of Nouns and very good examples of Verbs, but many subtly different sub-classes that fall somewhere in between. Each subclass possesses a "cluster" of grammatical properties that may or may not have any logical coherence. Consider an English word like *slurping*. Is this a noun or a verb? Well it can take a possessor, *his slurping bothers me*, but it doesn't easily occur with quantifiers, or certain other noun modifiers: ??*His many slurpings bother me*, ??*His much slurping bothers me*, ??*His fast slurping bothers me*. So *slurping* has some but not all properties of Nouns. On the other hand, it also has some properties of Verbs; it can take a direct object, as in *His slurping the soup bothers me*. It can also be modified with adverbial modifiers *His quickly slurping the soup bothers me*. Yet it doesn't take a nominative case subject, and cannot inflect like a Verb *He slurpings whenever he eats soup*. This particular cluster of properties cannot be attributed to some subclass, such as abstract nouns or nominalized verbs. For example, clear cases of abstract nouns cannot be modified by adverbials (*his truly sincerity*). Furthermore, some nominalized verbs with *-ing* take plurals more easily than others: *his many failings* vs. *?his many eatings*. This fact makes *failing* slightly more "nouny" than *eating*.

These subtle differences among the behaviors of various forms are probably not available to the fieldworker faced with thousands of forms, each potentially exhibiting a cluster of from zero to about 10 grammatical properties. To exhaustively categorize every word according to the particular cluster of Noun and Verb properties it exhibits is a potentially never-ending task, and may actually be distracting to readers. Therefore the concepts of Noun and Verb stand as imprecise approximations that nonetheless are *precise enough* to be useful in expressing important grammatical concepts (similar to the way *her hand* is precise enough a reference form in example 1). This is but one example of how being too inclusive can actually make a grammatical description less communicative.

4. Technical accuracy vs. Understandability

Accuracy is definitely a value in grammar writing. For this reason, many formalisms and abbreviatory systems have developed over the years as linguists have attempted to make their work as precise as possible. The problem is that formalisms (like language structures themselves) arise within particular communities, and are refined by generations of scholars and their graduate students in PhD dissertations, monographs, and research articles. Readers who lack a background in the specific analytic tradition employed by the grammar writer are likely to be mystified and put off by an over-reliance on formalism and theory-specific terminology. While formalisms and other analytic techniques may increase precision, they often do so at the expense of understandability of the text to future generations.

Furthermore, even the most elaborate mathematical formalisms are still not completely precise. As mentioned several times throughout this paper, language users employ conventionalized categories and patterns in all kinds of creative ways to communicate unique and nuanced ideas. Is there a "rule of English grammar" that can explain the structure of the following actual communicative exchange?

(2)   A. That boy is silly.
         B. He's not silly. He just be's silly when he's around girls.

Certainly speaker B (a 12 year old Anglo-American girl) had a rule in her grammar that made her response reasonable and communicative. And I venture to guess that most native English speakers will find B's response coherent and interpretable, even if they would never use it themselves (or at least would never admit to using it). But is it a rule of "English" (whatever that is), or simply a quirky "error" on the part of a less-than-fully-competent speaker? It is my contention that bending conventionalized patterns, and employing them creatively in new and unusual ways is the normal way that people communicate with one another. This is not bizarre, exceptional or erroneous use of language, as studies in corpus linguistics are beginning to show us.

If a grammar writer thinks that all such creative usages need to be incorporated into the written grammar, the task will never be complete. In the case of languages that lack a written tradition, it is especially difficult to determine which usages that appear in natural text are part of "the Grammar of the Language" and which can be chalked up to individual creativity. Of course the distinction is not absolutely clear, and even fully competent native speakers will not necessarily agree. Therefore, it

becomes another judgment call (a call for "balance") on the part of the writer as to how "accurate" one should be about describing the patterns of usages in natural texts.

5. Universality vs. specificity

Each language exhibits features common to all or many other languages, as well as features unique to that particular language. While grammar writers want their work to be usable and understandable by linguists working in other language traditions, and those studying universal characteristics of Language, they also want to highlight the unique and wonderful characteristics of the particular language they have spent so much time learning and analyzing. Often the concepts and terminology that have arisen in other language traditions do not seem to match the categories of the language being described very well, and so one is tempted to devise new and unique terms to describe these new and unique categories. Of course, the more new concepts and terminology are introduced into the written grammar, the more difficult it becomes for readers from other traditions to appreciate.

On the other hand, a language may exhibit a feature that is so distinct from what has been described in previous literature that a new term is necessary. If this is the case, the grammar writer must take care to define the new concept very carefully, and highlight the fact that this is truly new knowledge. For example, at present Doris Payne is grappling with the issue of how to label two tonally marked "case forms" in Maasai. Let's call them "Form A" and "Form B" for now. Form A is the citation form for nouns, and occurs when a noun (of any grammatical relation) occurs before the verb (examples 3a, d, and e) or when an Object noun comes after the verb (3c and d). Form B is used for Subject nouns that come after the verb (whether they are subjects of transitive or intransitive clauses, ex. 3b, c and e):

(3)   a. Ɔl-múrránì   o-ipid-ó.   'The warrior (FORM A) jumped.'
           MSG-warrior.FORM.A   3-jump-PF

        b. É-ípíd-ó   ɔl-mʊrraní.   'The warrior (FORM B) jumped.'
           3-jump-PF   MSG-warrior.FORM.B

        c. É-tóósh-ó ɔl-mʊrraní   ɔl-ásʊráí.   'The warrior (B) hit the snake (A).'
           3-hit-PF   MSG-warrior.FORM.B   MSG-snake.FORM.A

        d. Ɔl-múrránì   o-toosh-ó ɔl-ásʊráí.   'The warrior (A) hit the snake (A).'
           MSG-warrior.FORM.A   3-hit-PF   MSG-snake.FORM.A

        e. Ɔl-ásʊráí   é-tóósh-ó ɔl-mʊrraní.   'The warrior (B) hit the snake (A).'
           MSG-snake.FORM.A 3-hit-PF   MSG-warrior.FORM.B

The question is, how do we label these case forms in a way that genuinely helps readers who work in other language traditions understand the forms, while at the same time highlighting the special characteristics of Maasai? Some options one might consider include:

	Form A	Form B
1.	Object	Subject
2.	Accusative	(Marked) Nominative
3.	Absolutive	(Marked) Nominative
4.	Absolutive	Subjective
5.	Form A	Form B

Options 1-4 are all based on terminology from other established areal traditions, and all can be misleading for one reason or another. Option 1 (Object, Subject) is misleading, since Form A marks Subjects when they appear before the predicate (and in certain other contexts, such as predicate nominals). Option 2 is similarly misleading. Options 3 and 4, are also misleading in that the term "Absolutive" is usually employed in opposition to the term "Ergative." However, in Maasai, there is clearly no Ergative case. The term "Absolutive" in these options makes reference to the fact that "Form A" is the "naming form," or "citation form," i.e., the form that speakers naturally revert to when a noun occurs outside of any grammatical context.

Option 5 is somewhat of a cop-out, since the labels "Form A" and "Form B" make no reference whatsoever to familiar linguistic categories. Such terms are occasionally called for, when categories are so unusual that entirely new terms are needed to refer to them. Once a good friend of mine, David Watters, did a discourse study of two verb forms in Kham in which he underlined all instances of one form with a red pencil and all instances of the other with a blue pencil. After dealing with this analytic technique for some time, he found himself naturally referring to "red verbs" vs. "blue verbs", and he developed a rather sophisticated analysis of the functions of these forms in texts. Since the uses of these two forms did not seem to correspond to any previously established categories in the general linguistics literature, he continued to use the terms "red verb" and "blue verb" to gloss and refer to the two forms in his initial write-ups. This solution "worked" for David because he was able to provide content for otherwise grammatically meaningless terms in the process of doing his analysis. Eventually, after studying linguistic work in related languages, David found more "linguisticky" labels for these forms -- "conjunct" and "disjunct". These communicated well enough for those familiar with the literature on Tibeto-Burman languages, but still had to be explained in detail for general readers. Many such terminological quirks have resulted from similar decisions that linguists have made in the process of developing grammatical terminology for particular languages. Some examples that come to mind include "heavy" vs. "light" vowels, "strong" vs. "weak" conjugations, "soft" vs. "hard" consonants, and so forth.

6. Responsibility to the academic community vs. responsibility to speakers

There are many audiences or "stakeholders" in a grammatical description. The grammar writer must be able to balance the needs of all the various interested parties

in a way that acknowledges legitimate interests, even when they may seem to conflict with one another. Perhaps the major tension felt by grammar writers in this regard is their responsibility to their own career path, normally via recognition by authorities in their academic disciplines and universities, vs. their responsibility to the community of speakers that have made the research possible. Kadanya (2007) is an eloquent elaboration of this tension from the perspective of a speaker of an underdescribed endangered language (Toposa, a Nilo-Saharan language of Sudan).

There is an expanding literature at present on the ethics of linguistic fieldwork (see, e.g., Ameka, Evans and Dench 2006, Crowley 2007). Briefly, some of the issues this literature raises include:

- Informed consent. It is very important that speakers who provide the data that appear in a grammatical description understand the project, and the uses to which the data will be put. This is not always a straightforward matter. Universities sometimes require an "informed consent" form be filled out by every subject. However, in some communities a piece of paper with writing on it in any language is at best meaningless, and at worst a cause for suspicion. Other times, even careful explanations may not serve to convince speakers of the legitimacy and usefulness of the research. If the researcher obviously has access to money and power that is not available to speakers, speakers may legitimately suspect the researcher of "mining" the community for some valuable commodity that members of the community can only speculate on. In such cases, explanations couched in terms of "language documentation", "language preservation," "service to future generations", etc. are likely to be received with skepticism. Linguistic research is a long-term effort. Unlike community development projects such as wells, vocational training programs or electrification projects, the benefits of linguistic research only become apparent after many years. It may take weeks, months or even years of work before a linguistic fieldworker develops enough credibility to be accepted by a community.

- A commitment not to cause harm. How can linguistic fieldwork harm an individual or language group? Several ways are documented in Crowley (2007: 25-27). To a linguist, the content of example words, sentences and texts is secondary to the grammatical patterns that they illustrate. However, to speakers of a language, the ideas expressed are much more important than the forms used to express them. For example, a linguist doesn't care if a sentence like "John loves Mary" is true or not in some context, but only whether it is grammatically well-formed. However, if community members are familiar with a particular "John" and a particular "Mary," such an utterance being printed in a book or recorded electronically could raise a scandal and cause all kinds of problems for the named individuals or the community itself. For another example, Crowley (2007) documents cases where particular stories constitute evidence in land claim disputes. Publishing one clan's version of a story, rather than another's, may stir up long standing political and social tensions. Linguistic research involves collecting and publishing information, and information is power. One must constantly be aware of the potential problems created by the content of the data represented in a written grammar.

- Adequate compensation. People who work on a linguistic project are entitled to compensation for their time. Every hour spent working over linguistic data is an hour not spent in some other important activity, and for most people in the world life is a full-time job. Salaries paid to language consultants should be comparable to what individuals with similar levels of education would receive in the community where the work is being carried out. To pay more risks the possibility of creating an overdependence on the researcher, which in turn may complicate eventual community "ownership" of the project. When and if the project is adopted by a local organization, that organization will be saddled with the expectation that salaries paid to staff will be commensurate with salaries paid by the fieldworker. It is important to keep the project as "lean" as possible from the very beginning, with a view to turning it over to local control at some point in the future.

- Meeting felt needs. Speakers of underdocumented languages rarely feel a need for language documentation. Felt needs tend to be much more concrete and short-term, e.g., economic, healthcare and infrastructure development. The major contribution that linguists can make to a community's felt needs is in the area of education. Although there are exceptions, most people in the world value literacy and wider education. If a community lacks adequate educational opportunities, and if the researcher can make a credible case that linguistic research will help solve this problem, this can contribute to a community's enthusiasm for a linguistic project. Small tangible results such as calendars, topical dictionaries (bird names, plant names, etc.), story booklets, and health manuals can be encouraging, both to speakers and linguists alike. It is very important that such concrete results, with tangible value be produced early in the research program in order to help people "catch the vision" for what the linguist is doing and to encourage and maintain interest in the linguistic project.

- Making research results available to the community. Individuals who participate in linguistic research are entitled to access to the results of the research. This can be problematic for the researcher if, for example, the main result of the project is a doctoral dissertation or other treatise written for a particular academic committee or network. Such a treatise may need to be translated, not only into a language that community members understand, but also into a style of discourse that can be appreciated by a very different audience. That may mean writing a second pedagogical grammar, or "school grammar," with speakers of the language in mind as the audience. This can be a major effort, comparable to that of writing the academic treatise itself, and so fieldworkers may be tempted to skimp on this part of the project. However, since the research would not be possible without the cooperation of the community, logic and ethics require that the community understand and have access to the results of the research in a usable form. The main way to accomplish this is for speakers of the language, rather than the researcher, to coordinate the creation of materials destined to be used by the community. Not only does this relieve the field linguist of some of the work involved in producing this type of material, but also helps ensure that the result will genuinely be of interest and use to speakers. Mithun (2007) provides an excellent description of such a collaborative grammar writing project.

- Ongoing relationships. An academic researcher who comes to a community for a summer or a year, gathers data, and then leaves, never to be seen again, often does a disservice to the community who provides the data, and to future researchers. Unfortunately, there is a somewhat justified stereotype of outsiders expressing interest in (or, less charitably, "gawking at") speakers of underdocumented languages, who may seem "exotic" from the point of view of much of the world outside the community itself. Many communities around the world have become jaded and even hostile toward any kind of "research project" that may come along because of earlier negative experiences with outsiders. Researchers who enter a community with an "agenda" of their own, without being careful to understand the needs and show respect for the cultural mores of the people with whom they work, alienate the community and set a terrible precedent for others who may come in the future. As mentioned above, linguistic work is a long-term process. Relationship building is also long-term. A project that is properly integrated with the community is more likely to continue providing benefits after the researcher has left, than one that is imposed from the outside. One way linguistic fieldworkers can build and maintain such long-term relationships is to provide training, e.g., typing, computer skills, literacy, secondary education, or even higher education (in linguistics!) for individuals involved in the project. Facilitating training of community members to take over further language development work is probably the best way to ensure that the project will outlast the presence of an outside linguist.

Most of the points mentioned above apply most obviously to situations in which a field linguist is doing academic linguistic research in a community that is not the linguist's own. Increasingly, language development projects are being undertaken by linguists who are speakers of the language being documented. The above considerations apply in those situations as well, to varying degrees. Communities whose languages are endangered are beginning to value linguistic training and methodologies in order to foster and maintain language use and awareness. Communities whose heritage language is no longer in use turn to linguists, either members of the community or outsiders, to help preserve what records there may be of the language, and maintain a sense of language identity in the face of social and cultural pressure from dominant languages.

7. A form-driven vs. function-driven outline

Every language is a formal, structural system that arises in a human community in response to communicative needs. Every structural piece of a language has both a formal and a functional dimension. Thus, a grammatical description may be organized according to forms -- giving the function or functions of each form in turn --, or it may be organized according to functions -- giving the form or forms used to accomplish each function in turn. The outlines of most grammatical descriptions to date seem to combine these two perspectives to one degree or another, but largely without a principled reason for the division. Many problems arise when grammar writers fail to clearly identify forms and functions independently of one another, and therefore mix form-driven and function-driven description in a haphazard manner. For this reason, another area that requires careful balance in grammatical description is between form-driven and function-driven components.

The formal and functional dimensions of linguistic units are closely linked, but need to be identified independently of one another. For example the tool we call a "screwdriver" is named for one of its functions -- driving screws. What it IS is a thing designed specifically for driving screws. However, driving screws is only one of its possible functions. It can also be used for opening paint cans, scraping dirt out of tight corners, as a pointer in an academic lecture, or any number of other functions. It is a thing with a particular structure designed, or uniquely adapted, to serve a particular function. However, that same structure is also useful for other functions as well. The form of a screwdriver is thus logically distinct from any particular function that it might fulfill. It does not cease to be a screwdriver and suddenly become a can opener just because someone uses it to open a can.

Something similar is true of linguistic structures. For a simple example, consider the following expression:

(4)   the boy who puts them in his basket

This seems like a noun phrase modified by a relative clause. Most grammar books describe a relative clause is a clause that modifies a noun, and indeed that is probably the major function of most structures that are called relative clauses in the linguistics literature (see, e.g., Keenan 1985, Payne 1997 and many others). However, this structure can also serve quite a different function, as in the following extended example:

(5)   There's a man in the tree picking pears, and a boy on the ground with a basket. The man throws the pears down to the boy *who puts them in his basket*.

In this example the "relative clause" *who puts them in his basket* cannot be said to "modify" the head noun, *boy*, either restrictively or non-restrictively. Rather, this clause actually asserts a sequential event. First the man throws the pears to the boy, then the boy puts them in his basket. These are two "foregrounded" events in the event structure of the text (according to, e.g., Hopper and Thompson 1980). So this is one case (out of dozens or hundreds that could be provided) of how structures that primarily serve one function can be used to fulfill other functions.

7.1. Advantages of a form-driven description

Most of the grammars of underdocumented languages that have appeared in the last hundred years or so have been primarily form-driven. There are many good reasons why this is the case, including the following:

- A form-driven description is relatively easy to outline. Linguistic structures tend to be more categorical than linguistic functions. That is, form "discretizes" (makes into distinct categories) open-ended functional "space." For this reason, it is easier (though not a simple matter by any means) to identify particular forms, and situate them in an outline of a grammar, than it is to do the same with functions.

- A form-driven description is consistent with the way many students and teachers view "grammar" – a list of structural facts expressed as rules. Perhaps unfortunately, traditional approaches to first and second language teaching and linguistics have evolved with an emphasis on structures, often to the exclusion of functions. Structures, such as nouns, verbs, phrases, clauses, etc. are familiar (if not beloved) to students and teachers, and therefore a structure-driven outline "resounds" with expectations of what a grammar "should be."

- A form-driven description can be very clear. It is relatively easy to identify structures – identifying functions is more challenging. This point is closely related to the other two. Once you have a form-driven outline, the task of actually writing the grammar becomes a matter of "marching through" the outline. Each structural topic can be treated in its own autonomous section, and need not necessarily be influenced by other sections.

7.2. Disadvantages of a form-driven description

In addition to the advantages listed above, there are several disadvantages to a strictly form-driven outline, including the following:

- A form-driven description can be boring. Why is it that everyone (that is, normal people rather than linguists or grammarians) seems to hate grammar? Grammar is what allows people to communicate with one another, and everyone loves to communicate! One reason for this strange phenomenon, I believe, is the way grammar is conceptualized and taught (see above). Somehow we have gotten the idea that "grammar" consists of a list of impenetrable formal rules that must be memorized by rote. It is a logical "machine" consisting of structural parts that have no necessary relation to real life. This conceptualization has worked its way into first and second language grammar classes and even into the linguistics literature. A grammatical description based on this "grammar as machine" metaphor tends to be dry, boring and difficult to relate to the concerns of real people because it fails to take into account the fact that a written grammar is an act of communication. Payne (2007) is largely an argument against this conceptualization.

- A form-driven description emphasizes idiosyncratic facts of the individual language, making it more difficult to relate the language to typologically very different languages. Languages are similar in their functions, but quite different in the structures they employ to accomplish those functions. A strictly form-based description need not relate the language being described to other languages at all, because the universal need to communicate is not in focus. With the rise of typological linguistics, universal properties of human languages have become more of a focus than the sometimes idiosyncratic structural facts of particular languages.

- A form-driven description can misrepresent or fail to represent "functional systems" that span more than one word class or level of structure. Even as a form-driven description can obscure similarities from one language to the next, it can also obscure functional systems within one language. For example, the "tense system" of English spans at least two levels of grammatical structure:

(6)   You mock my pain!   Present   zero
        You mocked my pain!   Past   morphological
        You will mock my pain!   Future   analytic

In a form-driven grammar, present and past tense would be treated in the word-level chapter while future tense would be treated in the phrase level or syntax chapter. Thus the notion of a unified "tense system" would be lost.

7.3. Advantages of a function-driven description.

The advantages and disadvantages of a function-driven grammatical description are largely the mirror images of the disadvantages and advantages of a form-driven description described above. Nevertheless, I will enumerate these briefly in this subsection and the following:

- A function driven description acknowledges the common sense fact that language serves a purpose – namely communication.

- A function driven description brings together different structural pieces that conspire to accomplish ranges of communicative functions ("functional systems").

- A function driven description makes comparison among typologically distinct languages more possible.

7.4. Disadvantages of a function-driven description.

- A function driven description can be hard to outline. Functions are not discrete and categorical, and therefore it is challenging to identify them, and organize them into a coherent outline.

- A function driven description tends to make typologically distinct languages seem more similar to each other.

- A function driven description can be "Open ended" in that almost anything can serve almost any function, given enough context. For a simple example, a "passive" construction is often defined functionally as one that "downplays" an AGENT and "upgrades" a PATIENT. Well if that is the definition of passives, then would the following English sentences all be passives?

(7)   The glass broke. (Downplaying the AGENT who broke the glass.)
        Some guy broke the most beautiful vase in the world.
        These jeans wear easily.
        John underwent surgery.
        As for okra, I can't stand it.
        Okra is what I can't stand.

These all can be construed as somehow "downplaying an AGENT" and/or "upgrading a PATIENT." Without clear structural guidelines, there is no way to constrain the range of sentences that might be construed as fulfilling a given

function. For this reason, I would like to argue for a balanced formal/functional approach to linguistic description, to which I turn in the next section.

7.5. The solution: A balanced formal-functional approach

The type of grammatical description that I would like to recommend is one that employs a form-driven approach for those areas of grammar that are the most controlled, systematic and rule-dominated, and a function-first approach for those areas that tend to cross-cut structural levels. The controlled, systematic and rule-dominated parts of language include:

- Phonology (excluding intonation).

- Morphophonemics

- Inventory of derivational morphology (which derivational categories apply to which roots, etc.)

- Inflectional inventory (determining the range of inflectional possibilities for person and number "agreement" and case marking)

- Pronoun inventory (isolating the entire set of free pronouns or pronominal clitics)

- Lexical inventory (acquiring the words for a large number of culturally significant things and activities)

Notice that in this section there is an emphasis on obtaining *inventories* of various forms. In many cases, this kind of information is best obtained via direct elicitation. This is because languages typically employ a small number of forms in text, though many more forms are logically possible. Full paradigms are seldom constructable based on data that appear in natural texts alone. For example, a declarative sentence with a second person subject is very rare in texts, because people don't often inform other people concerning activities of the person spoken to, e.g., *You are baking bread*. Questions are much more natural in such a context. Nevertheless, a description of the language would be incomplete if the second person declarative forms were missing. Elicitation is essential to the completion of paradigm charts.

Often the meaning of a particular morpheme or construction is not clear until the entire range of possibilities that could replace it is identified. The same observation can be applied to syntactic constructions. For example, whether a particular transitive construction is a passive or an ergative depends at least partially on whether there exists a corresponding "active" construction. Similarly, the precise function of Subject-verb-Object word order may not be apparent until minimal pairs with Verb-subject-object order are obtained. Text data may exhibit other orders, but in examples extracted from texts, there are usually enough other formal differences that the precise contribution of word order to the observed semantic differences is obscured. True minimal pairs are usually obtainable only through elicitation.

The more pragmatic, semantic and subtle parts of language are best approached and analyzed from a function-first perspective, via a large body of naturally occurring text, supplemented by elicitation where necessary. These would include:

- Intonation.

- Constituent order.

- Inflectional morphology (determining the precise functions, including tense/aspect/mode).

- Voice (alignment of grammatical relations and semantic roles of verbal arguments)

- Sentence level particles (evidentials, validationals and pragmatic highlighting particles).

- Clause combining (including relativization, complementation, adverbial clauses and clause chaining)

- Lexical semantics (determining the nuances associated with various lexical choices, including derivational morphology and pronouns).

- Pragmatically marked structures, such as clefts, questions, etc.

8. Conclusion

Field linguists have several goals in mind when approaching the task of writing a reference grammar. These include:

- Communicativity. A grammar should clearly communicate complex facts.

- Inclusivity. A grammar should describe ALL the grammatical features of a language.

- Usability. A grammar should have a well-defined audience who will find the grammar of genuine use.

- Accuracy. A grammar should be as technically accurate as possible.

- Universality. A grammar should relate the language described to known universal principles of human language.

- Specificity. A grammar should highlight the unique and beautiful features of the language described.

- Integrity. A grammar should acknowledge and fulfill the fieldworker's ethical responsibilities to the speech community.

Unfortunately, these goals often seem to conflict with one another, and so grammar writers must sometimes partially compromise one goal in order to fulfill another. Many potential grammar writers are paralyzed by tensions caused by such conflicting goals, as they approach the complex task of writing a linguistic grammar. In this paper I have described several of these tensions, and have argued for a sense of "balance" in grammatical description. I hope that the suggestions made in this paper will encourage writers of descriptive grammars by acknowledging the tensions, and giving grammar writers ways to balance competing goals. Grammatical description is an art as well as a science. Writers of linguistic grammars must remember that they are artists creating an esthetically pleasing and engaging piece of non-fiction literature, as well as scientists producing a precise and informative research report.

References

Ameka, Felix, Alan Dench, and Nicholas Evans (eds.). 2006. *Catching language: The standing challenge of grammar writing.* Berlin: Mouton de Gruyter.

Becker, Alton L. 1979. Text-building, epistemology, and aesthetics in Javanese shadow theatre, in *The imagination of reality: essays in Southeast Asian coherence systems* ed. by A. L. Becker and Aram A. Yengoyan, 211-243. Norwood, N. J.: Ablex.

Crowley, Terry. 2007. *Field linguistics: a beginner's guide*. Oxford University Press.

Grice, H. Paul. 1975. Logic and conversation, in *Speech Acts* [*Syntax and Semantics* vol. 3], ed. by Peter Cole and Jerry L. Morgan, 41-58. New York: Academic Press.

Hopper, Paul and Sandra A. Thompson. 1980. Transitivity in grammar and discourse. *Language* 56, pp. 251-99.

Kadanya, James Lokuda. 2007. Writing grammars for the community, in *Perspectives on grammar writing* [*Benjamins Current Topics* 11] , ed. by Thomas E. Payne and David J. Weber, pp. 19-23. Amsterdam and Philadelphia: John Benjamins.

Keenan, Edward L. 1985. Relative clauses, in *Language typology and syntactic description, vol II: complex constructions*, ed. by Timothy Shopen, pp. 141-70. Cambridge University Press.

Mithun, Marianne. 2007. Grammars and the community, in *Perspectives on grammar writing* [*Benjamins Current Topics* 11] , ed. by Thomas E. Payne and David J. Weber, pp. 45-69. Amsterdam and Philadelphia: John Benjamins.

Payne, Thomas E. 1997. *Describing morphosyntax: a guide for field linguists.* Cambridge University Press.

Payne, Thomas E. 2007. A grammar as a communicative act, or what does a grammatical description really describe? in *Perspectives on grammar writing* [*Benjamins Current Topics* 11] , ed. by Thomas E. Payne and David J. Weber, pp. 127-142. Amsterdam and Philadelphia: John Benjamins.

Sperber, Dan & Dierdre Wilson. 1995. *Relevance: Communication and cognition* (2nd Edition). Oxford. Basil Blackwell Ltd.

Appendix: A possible outline for a balanced grammatical description

The following is one possible outline for a grammatical description that tries to "balance" form-driven vs. function-driven approaches. This outline is humbly offered as a source of ideas and guidance to fieldworkers who feel a need for such guidance. As discussed in the paper itself, every grammar exhibits unique characteristics based on the interests, goals, abilities, and personality of the grammar writer, as well as characteristics of the language and of its sociolinguistic situation. Hence this is not meant to be a "checklist" a "field manual" or a "strait-jacket," but simply a source of ideas for elaborating a grammatical description.

Items followed by an asterisk (*) are considered essential. Other items may or may not appear in the grammar outline, depending on a) the intended use of the grammar, b) the special experience and interests of the author and c) the individual characteristics of the language. Of course any particular grammar may also include more headings than what are found here.

Front matter

Acknowledgements (*)

Introduction (Including theoretical assumptions and purpose of the grammar.)

List of abbreviations (*)

Part I: The Cultural, Ecological and Sociolinguistic Context of the Language

1.1 The name of the language (*) Including "endoethnonyms" (the name people use to refer to themselves) "exoethnonyms" (terms used by outsiders to refer to a particular ethnic group).

1.2 Previous research (*)

1.3 Demography (*) Number of speakers, location and other linguistic groups in the area.

    1.3.1 Map(s)

    1.3.2 History/migrations

1.4 Ecology

1.5 Ethnography (material culture, cosmology) (*)

1.6 Genetic and areal affiliations (*)

1.7 Literary traditions

1.8 Dialects (including classical/written varieties if applicable) (*)

1.9 Sociolinguistic situation (*)

    1.9.1 Multilingualism and language attitudes

1.9.2 Contexts of use and language choice (*)

1.9.3 Viability (*)

1.9.4 Loan words

1.10 The corpus (*)

1.10.1 The nature of the research (affiliation, location, duration) (*)

1.10.2 Consultants and other sources (*)

1.10.3 Presentation of data (*)

Part II: Structural Overview (form driven approach)

2.1 Typological Sketch

2.2 Phonological inventory and orthography (*)

2.2.1 Consonants (*)

2.2.2 Vowels (*)

2.2.3 Tone / stress (*)

2.3 Phonetics (*)

2.4 Syllable structure

2.5 Word structure

2.6 Major phonological and morphophonemic processes (*)

2.6.1 Process 1 (*)

2.6.2 Process 2 (*)

2.6.3 Process 3

2.6.4 Process 4

2.7 Relaxed speech rules and contractions

2.8 Word Classes (*)

2.8.1 Nouns (*)

The structure of the noun word

Derivational processes

Inflectional processes

Count vs. mass nouns

Proper names

Other grammatically distinct subclasses of nouns

2.8.2 Pronouns and/or anaphoric clitics (*)

Personal pronouns (*)

Demonstrative pronouns (*)

Other (relative pronouns and question words may be more efficiently treated   in the sections on relative clauses and questions.)

2.8.3 Verbs (*)

Verb structure (a diagram of a verb and its morphology) (*)

Derivational (stem-forming) processes

Inflectional processes

Grammatically distinct verb subclass 1

Grammatically distinct verb subclass 2

Grammatically distinct verb subclass 3

Grammatically distinct verb subclass 4

2.8.4 Modifiers

Descriptive adjectives

Non-numeral quantifiers

Numerals

Adverbs

2.8.5 Auxiliaries

2.8.6 Ad-positions (prepositions or post-positions)

2.8.7 Particles or other minor word classes

2.9 Constituent Order Typology

2.9.1 Constituent order in main clauses

2.9.2 Constituent order in verb phrases

3.5.1 Constituent order variation

3.5.2 Contrastive/emphatic particles

3.5.3 Contrastive/emphatic intonation patterns

3.5.4 Cleft constructions

3.5.5 Negation (*)

3.5.6 Questions (*)

    Yes/No Questions (*)

    Question word (information, content) questions (*)

3.5.7 Imperatives (*) (including subtypes)

3.6 Clause Combinations (*)

3.6.1 Serial verbs

3.6.2 Complement clauses (*)

3.6.3 Adverbial clauses (*)

3.6.4 Clause chaining, medial clauses and switch reference

3.6.5 Relative clauses (*)

3.6.6 Coordination

3.7 The Language in use

3.7.1 Lexical Typology

    Space, direction and motion

    Causation

    Perspective

    Salience of semantic features

3.7.2 Continuity (cohesion) and discontinuity

    Topic (referential) continuity

    Thematic continuity

    Action continuity

3.7.3 Episodic prominence

Climax/peak

Intensification

3.7.4 Genres

Conversation

Narrative

Personal experience

Historical

Folk stories

Mythology

Hortatory

Procedural

Expository

Descriptive

Ritual speech

3.7.5 Miscellaneous and conclusions

Idiomatic expressions / proverbs

Sound symbolism

3.7.6 Summary of typological findings

End matter

Text with interlinear translation
Glossary
References cited (*)
Index (*)