

Grammar writing from a dissertation advisor's perspective

Andrew Pawley
Australian National University

Anyone who intends to produce a grammar of a previously little-described language needs to (1) plan the scope, methods and timetable of the data gathering process, (2) think about the conceptual framework that will shape data-gathering and analysis, (3) gather and organize the data, (4) analyse the data, and (5) plan the structure of the written account and (6) write the grammar. The steps are not simply sequential but are to some extent cyclical. This chapter will look at an advisor's role in guiding a PhD student through these steps. It will focus on the following questions: What kinds of data, and how much, are sufficient to base a grammar on? What is a realistic size for a PhD dissertation grammar? What are the main alternative ways of organizing a grammatical description, e.g. in terms of topic divisions and sequencing? What are the *dos* and *don'ts* to be followed in order to make the grammar as descriptively adequate and user friendly as possible? What are the main reasons why some students take forever to complete the analysis and writing process?

1. AIMS. This chapter looks at grammar writing from the perspective of a dissertation advisor. It identifies key decisions that must be made in the course of planning and producing a grammatical description for a doctoral dissertation, and comments on the role of an advisor in dealing with these matters.¹

It is more than 40 years since I was a PhD student myself, doing fieldwork on Kalam, a Papuan language of New Guinea, and writing a dissertation on the phonology and grammar of that language under the supervision of Bruce Biggs at the University of Auckland. Following those *rites de passage* I got a job as a linguist, and since then have in my turn supervised more than 50 graduate theses, about half of them reference grammars. But I am very mindful of the fact that there are others who have written quite extensively on aspects of the craft of grammar-writing, whereas I have not done so until now.² The remarks that follow will draw on the observations of many other scholars as well as on my own experience.

I will consider the following general questions to do with writing a grammar for a dissertation project.

- What is the nature of the task? What is entailed in writing a grammar?
- How should one prepare for such a project?
- What are the main decisions to be made on the way?
- What are the *dos* and *don'ts* in presenting the analysis?
- How not to write a grammar. Some bad strategies in data collecting and writing up.
- What is the role of the advisor in these matters?

¹ I am indebted to the two anonymous referees for helpful comments on a draft.

² For example, Ameka et al. (2006), Bowerman (2008), Mosel (2011), Payne (1997) and Payne and Weber (2005), to name just a few.

2. WRITING A GRAMMAR: WHAT IS THE NATURE OF THE TASK AND HOW SHOULD ONE PREPARE FOR IT?

2.1 THE MAIN STEPS. The expression 'to write a grammar' has a number of different readings. A literal reading is roughly 'to give a written account of the grammar of a language' but as a description of what is entailed in producing a grammar this is misleading. The writing bit is just one of several parts of the job. Anyone who would produce a grammar must go through at least the following stages: (1) plan the scope, methods and timetable of the data-gathering process, (2) think about the conceptual framework that will shape data-gathering and analysis, (3) gather the data, (4) organize and analyse the data, (5) plan the structure of the written account and (6) write the grammar.

However, we should bear in mind that the steps are not just sequential but also cyclical to some extent. One begins analysis on day 1 of the data-gathering phase and preliminary analyses will influence or direct the next stage of data gathering. The process of writing up will reveal gaps in the data or flaws of analysis, leading to further data-gathering, reanalysis, and a revised presentation, and so on.

2.2 PLANNING THE PROJECT. Many things have to be planned at the outset, beginning with the choice of a language. One needs to plan the scope and timetable of the research, such as the number and length of field trips and the methods of data collection, and one must give some thought to the model of grammar that will provide the framework for the analysis and write-up.

2.2.1 WHAT KIND OF GRAMMAR? The primary readership of a PhD grammar will be professional linguists. The utilitarian purpose of a PhD dissertation is to demonstrate that the author has achieved a professional level of competence in the field of study and has made a substantial original contribution to it.³ Two things follow from these circumstances. First, the grammar should be an analytic grammar, a reference grammar, rather than a pedagogical one. Second, it will usually be of a relatively little-described language, at least one for which there exists no good reference grammar.

What about languages for which quite good reference grammar or grammars already exist, e.g. Indonesian, Maori, Mohawk or Quechua? Are there circumstances in which it is acceptable to do a grammar of such a language for a dissertation? Some advisors would say, yes, if there are data that cry out for a fresh analysis, or if the grammarian adopts a particular theoretical framework that one can be confident will reveal new insights. The risks in such cases are that the outcome will end up not being very novel. For reasons of national or academic politics sometimes students choose to describe a prestigious language that is already well documented rather than taking on a minority language that is undescribed but considered less important. I was told that this happened in the 1950s when several

3 While all linguists recognise that writing a descriptive grammar is a very challenging exercise, some would raise practical and/or ideological objections to writing a grammar as a PhD assignment. Instead, they would encourage students to do a dissertation that tackles a particular theoretical problem, or at least to write a theory-focused grammar, one that tests a particular theoretical framework. This, it is argued, will give them a better chance of getting a job in a field where theoretical work is more highly valued than descriptive work.

Indonesian students studying at US universities each chose to do descriptions of Indonesian for their dissertation. In Indonesia at that time there was no career advantage to be gained from choosing to work on a minor language.

A compromise is to do a 'topics in the grammar of X' dissertation, rather than a general reference grammar, where the topics are just those domains where the analyst has an original contribution to make.

2.2.2 CHOOSING A DESCRIPTIVE FRAMEWORK. In principle, a firm decision as to what kind of descriptive framework to use need not be made in the initial planning stages. However, there are good practical reasons why student and advisor should consider this matter early on and come to a joint understanding. For one thing, a student's choice of advisors may depend on decisions made. You don't want your chief advisor to quit in the middle of your writing-up phase because he or she objects to the theoretical framework you are using. Moreover, the kinds of data to be collected will, to some extent, be dependent on the questions that the framework focuses on.

If a reference grammar is to be readable generations from now it should use a descriptive framework (i.e. a body of analytic concepts and terms) that is familiar to most or all grammarians. This is easier said than done because all descriptions are to some extent theory-specific and specific theories of grammar are notorious for having a short shelf life – and even in their heyday are accessible only to specialists. After about 1960, the research agendas of descriptive and theoretical linguistics diverged sharply. The ultimate aim of theoreticians is to arrive at generalisations about human language and to explain why languages are the way they are and why languages change in the ways they do. Many theoretical linguists have little interest in describing particular languages, regarding them merely as a source of evidence for choosing between competing theories or competing claims about linguistic universals. Descriptive linguists, by contrast, generally maintain an interest in developments in theory as being relevant to the systematic analysis of a language.

Accordingly, most grammar writers of my acquaintance take the view that where possible one should try as far as possible to avoid analytic concepts that are local to particular models of grammar and use instead what R.M.W. Dixon and Matthew Dryer call 'basic linguistic theory' (BLT) (Dixon 2009, 2011, 2012, Dryer 2006). By this is meant, roughly, the analytic concepts that have widest currency among descriptive linguists and typologists. There is a very sizeable body of such concepts. Many are grounded in traditional grammar.

However, BLT is not a fixed thing. The tools of grammatical and phonological analysis continue to evolve and one should be ready to take on board new and useful concepts. Given that such change is inevitable and healthy, it is probably unrealistic to expect that any grammar written today will be an easy read in two centuries time. But by carefully defining key analytic concepts, a grammarian can at least make the path easier for readers.

In any case, the writer of a grammar should be eager to ask questions that have been generated by theory-specific work, if these offer promise of throwing light on the grammar of the target language. Keren Rice (2006:403) remarks that "... the grammar should be informed by theory. This will help make it coherent, and it will allow questions to be asked

that might not come up otherwise. But...theory is not the goal of a grammar". If you want to discuss theoretical issues in depth you write a separate paper or book.

But for some students, who have a keen interest in or allegiance to a particular theoretical model, it is hard to resist the temptation to use that model. And I know at least one linguistics professor who would only agree to chair a student's committee if the student writes the grammar using a particular preferred model of generative grammar with elaborate formalisms. The upside of this was that students gained a rigorous training in that model and, often, achieved some insightful analyses. The downside was that the resulting grammars were overloaded with theory-specific formalisms that are off-putting to all but the most dedicated reader.

2.2.3 ON FIELDWORK. If the data are to come mainly or wholly from fieldwork among speakers of a little-described language in a remote location, how long should a student spend in the field? The field worker should aim to obtain sufficient data to write a first grammar and to gain a pretty good level of competence in the language. Gaining such competence is advantageous for a variety of reasons. The more fluent you are in the language the easier it is to talk to a wide range of people in the community, to argue the point, to follow conversations and monologues, to ask complicated questions and to elicit suitable illustrative examples, and, importantly, the better placed you are to critically evaluate the information that comes in and so to reduce errors.

It is impossible to give a hard and fast recommendation about the time needed to achieve these objectives, because this will vary according to (a) the difficulty of the language, (b) whether or not there are good descriptions available for closely related languages, (c) how difficult fieldwork conditions are, and (d) personal factors. However, as a general rule, I think one needs at least nine or ten months, and preferably a year in the field.

How should the fieldwork be apportioned? How many field trips is an optimal number? How should they be spaced? There are various reasons why it is better to have at least two shorter field trips than a single long one. The first spell of fieldwork should be the longer one, of at least six months, aimed at yielding enough material to draft a substantial part of the dissertation. One can then return home, refresh the body and mind, write up first drafts of a large section of the dissertation, see where all the gaps in the data are, and after six months to a year, return to the field to check and fill in gaps. Nicholas Evans (pers. comm.) argues in favor of three trips: the first and last quite short, say three months each, with a longer middle trip. The reason for preferring a relatively short first trip is that during the early stages of research on an unfamiliar language the researcher needs to spend a lot of time on analysis, simply making sense of the data, and such analysis can be done at home.

Circumstances may sometimes limit the student to a single, long field trip. In these cases it is advantageous to press ahead with analysis and sketch the core chapters while still in the field, so that problems of analysis and major gaps in the data will become evident while there is still time to remedy them.

The situation is different if the field site is easily accessible. And obviously, it is different for those students who have had the advantage of having spent years studying the language before they began their PhD. Students who are members of the Summer Institute of Linguistics (SIL) often have this advantage.

3. SOME NOTES ON DATA-GATHERING

3.1 THE IMPORTANCE OF LEXICAL DATA. Keeping a lexical file is an important part of grammar-writing. The file should include notes on the grammar of each lexical item and examples of use. As Lichtenberk (2008:6) points out, "Grammatical rules, or patterns, are generalizations over various properties of individual lexical items. One cannot write a reasonably detailed grammar...without fairly extensive lexical information". When working on his grammar of Toqabaqita, he found grammar- and dictionary-writing to be mutually beneficial:

Detailed grammatical analysis enabled a more accurate treatment of the grammatical elements in the dictionary than would otherwise have been the case. And the lexicographical work has been of great importance to the grammatical analysis. In any language, grammatical rules, patterns, regularities are of highly different degrees of generality. Few, if any, hold across the board. Many grammatical patterns are lexically sensitive; they hold for some but not all members of a certain word class.

Of course, it is unreasonable to expect someone who is doing a PhD dissertation on the grammar of a language to try to compile a comprehensive dictionary at the same time. The aim should be to test the grammatical properties of a representative sample of lexical items, so that one can arrive at a reasonably fine-grained treatment of word classes and sub-classes. This brings us to the next question.

3.2 HOW MUCH DATA IS NEEDED TO WRITE A GRAMMAR? How much data is needed to write a reference grammar? Obviously, that depends on how comprehensive the grammar is. One can probably write quite a useful grammar sketch based on a corpus of 5,000 words but it will contain many glaring gaps. 20,000 words will yield a more complete analysis but there will still be many gaps in the kinds and details of constructions represented. 30,000 and 40,000 words will yield still more complete analyses, and so on. But we need to bear in mind that some morphological forms and other grammatical combinations will not occur in a corpus of a million or even 10 million words. Directed eliciting is needed to collect morphological and phrasal paradigms. And that brings us to another question.

3.3 WHAT KINDS OF DATA ARE LEGITIMATE? What kinds of data are legitimate for basing a grammar on? Is elicited data acceptable or should all data come from natural discourse? Is there a desirable balance between the two? And when there is significant variation relating to age, dialect, etc., how one should handle this in gathering samples of the language?

No one would deny the value of a large corpus of natural discourse data. One should collect and transcribe extensive text materials of various genres. But it would be foolish to take an extreme purist position and exclude elicited data. What you can collect in ten months will not be sufficient to answer many questions about grammar and lexicon. You need to elicit paradigms, word meanings and sentence forms and translations, and to train

consultants to give you grammaticality judgments. Of course one should be alert to the dangers of using elicited data and so should check and double check where possible.

The corpus of data obtained during the early stages of fieldwork, when the linguist does not yet know the language well, is likely to contain lots of mistakes and this is especially true of elicited data. I recall the comments of a colleague on the efforts of three linguists who had independently carried out dialect surveys of many Fijian languages and dialects, collecting, among other things, a 300 word basic vocabulary list from each, using Bauan, the lingua franca of Fiji, as the eliciting language. He compared the results from one particular language, collected (as far as I recall) at the village of Namuamua, in the Namosi region of the island of Viti Levu. He said that linguist A visited Namuamua for a single afternoon and his vocabulary list contained 27 errors. B stayed for two days and his list contained 9 errors. (I was linguist B.) C stayed a month and was confident that his list contained no errors. C, who was the commentator, had the additional advantage of having near native fluency in the lingua franca and an impressive familiarity with virtually all of the Fijian languages and dialects.

4. THE SCOPE OF THE GRAMMAR

4.1 HOW LARGE SHOULD A PHD GRAMMAR BE? One of the pitfalls in writing a dissertation is trying to do too much. My supervisor told me firmly: "Remember, the PhD is just a qualification, a ticket to a job. Do a good job but nothing too grand. Leave the grand projects till later". Mark Donohue recalls that when he proposed to add yet another chapter to his already long dissertation draft, he was told gently "You know Mark, you're allowed to do research after the PhD".

OK, but just how large and how comprehensive should a PhD reference grammar be? Let's start with size.

There are some very large reference grammars of previously little-described languages. Exceptional are Keren Rice's grammar of Slave, which runs to over 1400 pages (Rice 1989) and Frank Lichtenberk's grammar of Toqabaqita, 1375 pages (Lichtenberk 2008). These are comparable to the largest grammars of English, e.g. by Quirk et al. (1974, 1100 pages) and Huddleston and Pullum (2002, 1800 pages). My home library contains a few other very other large grammars of Pacific Island languages, e.g. Mosel and Hovdhaugen's (1992) grammar of Samoan is around 800 pages and Alexandra Aikhenvald's (2008) Manambu grammar is around 700. However, none of these works were PhD theses. All were done by established scholars, or teams of scholars, over many years. In the case of the English grammars, the authors could build on centuries of previous work.

What is a reasonable length to aim at for a PhD grammar? I compared 18 grammars published between 1994 and 2008, each grammar being a revised version of a PhD dissertation submitted either to an Australian (13) or a Dutch or German university (5).⁴ The languages described are Austronesian (10), Papuan (6) and Australian (2). The figures without parentheses represent the number of pages of the grammar proper. The figures in parentheses indicate the total pages including appendices (usually texts) and references. The year of publication is given; often this is a few years after the dissertation was completed.

⁴ I do not have on hand details for a comparable sample of grammars submitted as dissertations to North American universities during this time frame.

TABLE 1 Details of 18 PhD dissertation grammars published between 1994 and 2008

| AUTHOR | TITLE | NO. PAGES |
|--|--|-----------|
| Australian National University (10) | | |
| Bowden | <i>Taba: Description of a South Halmahera language</i> (2001) | 408 (451) |
| Bugenhagen | <i>A grammar of Mangap-Mbula, an Austronesian language of Papua New Guinea</i> (1995) | 355 (418) |
| Ezard | <i>An Austronesian language of the Milne Bay area, Papua New Guinea</i> (1997) | 297 (320) |
| Farr | <i>The interface between syntax and discourse in Korafe, a Papuan language of Papua New Guinea</i> | 417 (459) |
| Hyslop | <i>The Lolovoli dialect of the North-East Ambae language</i> (2001) | 438 (476) |
| Jones | <i>Towards a lexicon-grammar of Mekeo (an Austronesian language of western central Papua)</i> (1998) | 553 (601) |
| Obata | <i>A grammar of Bilua, a Papuan language of the Solomon Islands</i> (2003) | 281 (329) |
| Quick | <i>A grammar of the Pendau language of central Sulawesi, Indonesia</i> (2007) | 601 (716) |
| Teng | <i>A reference grammar of Puyuma, an Austronesian language of Taiwan</i> (2008) | 279 (309) |
| van Klinken | <i>A grammar of the Fehan dialect of Tetun</i> (1999) | 322 (355) |
| Other Australian (3) | | |
| Eades | <i>A grammar of Gayo, a language of Aceh, Sumatra</i> (2005) | 316 (350) |
| Pensalifini | <i>A grammar of Jingulu, an aboriginal language of the Northern Territory</i> (2003) | 240 (262) |
| Sharp | <i>Nyangumarta, a language of the Pilbara region of Western Australia</i> (2004) | 392 (429) |
| Netherlands (4) and Germany (1) | | |
| Dol | <i>A grammar of Maybrat, a language of the Bird's Head Peninsula, Papua Province, Indonesia</i> (2007) | 290 (328) |
| Klamer | <i>Kambera, a language of Eastern Indonesia</i> (1994) | 336 (368) |
| van den Heuvel | <i>Biak: description of an Austronesian language of Papua</i> (2006) | 423 (473) |
| van Staden | <i>Tidore, a linguistic description of a language of the North Moluccas</i> (2000) | 355 (566) |
| Wegener | <i>A grammar of Savosavo, a Papuan language of the Solomon Islands</i> (2008) | 330 (372) |

It can be seen that 16 of the grammars (excluding appendices and texts) fall between 240 and 440 pages, with just a couple of outliers in the 600 range. The median length is 336 pages. Appendices with texts usually amount to 30-50 pages but in one case (the Tidore grammar of van Staden (2000) they run to 200 pages. Over the years I have generally advised students that 300-350 pages, excluding appendices, should be ample. However, my impression is that in recent decades PhD grammars have grown larger. If so, no doubt one

factor driving it is competition: survival of the fittest in the job market, Darwinian natural selection.⁵ Another factor (see discussion below) is progress in the field: we know more than we used to about some aspects of grammar and pragmatics.

4.2 WHAT SHOULD BE IN A REFERENCE GRAMMAR? HOW SHOULD IT BE ORGANIZED? It is generally agreed that a grammar should include descriptions of the phonology, morphology, syntax and semantics of a language. However, within these broad limits grammars vary considerably in scope and degree of detail. My impression is that reference grammars today, including those done as PhD dissertations, cover a wider range of phenomena than those done, say, in the 1950s, 60s and 70s, particularly in the domains of syntax and pragmatics. It is also my impression that the modern grammars tend to be more discursive and readable than those produced at the height of the structural and transformational-generative grammar eras. In those times grammarians were more concerned with form than function. Descriptions usually said less than nowadays about the semantic characteristics of word classes and little about such things as pragmatics and information structure. Treatments of syntax were more limited than today. Grammatical relations like subject and direct object were not so carefully defined.

However, it is also my impression that modern descriptions tend to be less rigorous than grammars written by structuralists of 40 and 50 years ago in one respect: their treatment of the combinatorial possibilities of constituents. Even before Chomsky, structuralist grammars sought to specify all and only the possible combinations of elements within whatever units were being described. Modern grammars tend to contain fewer generative formulae but offer a more expansive and more readable account of grammatical constructions. I think there is room in a grammar for both approaches. One can always insert formulae predicting the possible combinations at the end of a more expansive discussion.

When it comes to the finer details of a description—which grammatical phenomena to treat and in how much detail and in what order, most of us would agree with those, like Rice (2006:400-1), who advises against following a predetermined outline because each language demands its own strategy. She illustrates by referring to several grammars that differ markedly in the way they are organized.

However, there are some favoured patterns. I compared the contents of 17 different grammars of Austronesian and Papuan languages which were originally submitted as PhD theses. They fell into two classes. Most consist of 7 to 10 chapters, some run to 14-16 chapters. I also looked at some larger grammars that are not PhD theses, which have between 20 and 40 chapters.

In the case of grammars with 7-10 chapters, most of the chapter titles in the different grammars are essentially the same and the order of chapters is very similar. A typical sequence is:

5 I am aware of one PhD grammar that is right off the scale: Alexandre François's description of Mwotlap, a language of north Vanuatu. This totals 1033 pages or about 470,000 words, excluding appendices. It was completed in four years including a year's fieldwork. During the same period the author managed to do the research for, write up and publish a 350 page grammar of another language of Vanuatu. But mere mortals should not attempt such feats.

1. Introduction. (Also called The language and its speakers, or similar.)
2. Phonetics and phonology.
3. Word classes. (Also called Parts of speech.) Sometimes includes phrase types.
6. Derivational morphology.
5. The noun phrase. (Also called Referring expressions.)
6. The verbal complex.
7. Basic clause structure.
8. Complex sentences.

The content and organization of the PhD grammars with 14 to 16 chapters differs from these in two ways: (i) They give separate chapter status to items such as pronouns, numerals and adpositional phrases, which in other grammars are treated within one of the standard chapters. (ii) They devote whole chapters to construction types that are particularly prominent or elaborate in the language, e.g. serial verb constructions, possessive constructions. Three examples follow:

Lolovoli, NE Ambae (Hyslop 2001)

1. Introduction. 2. Phonology. 3. Basic clause structure. 4. Word classes. 5. Noun phrase. 6. Adjuncts. 7. Possessive and associative constructions. 8. Spatial reference. 9. Verb phrase. 10. Serial verb constructions. 11. Valency change and rearrangement. 12. Reduplication. 13. Existential, equational and locational clauses. 14. Subordinate clauses. 15. Coordination.

Taba, S. Halmahera (Bowden 2001)

1. Introduction. 2. Phonetics and phonology. 3. Morphological and syntactic units. 4. Parts of speech. 5. Basic clause types. 6. An overview of clausal syntax. 7. Nouns and noun phrases. 8. Verb morphology and valence. 9. Possession and related constructions. 10. Quantifiers. 11. Demonstratives and directionals. 12. Serial verb constructions. 13. Adpositional phrases. 14. Clausal modifiers. 15. Questions and requests. 16. Interclausal relations.

Fehun Tetun, East Timor (van Klinken 1999)

1. Introduction. 2. Phonology and morphophonemics. 3. Word classes. 4. Derivational morphology. 5. Numerals, classifiers and the numeral phrase. 6. Pronouns and determiners. 7. The noun phrase. 8. Prepositions and the prepositional phrase. 9. The clause. 10. Auxiliaries. 11. Adverbs and verbal modifiers. 12. Serial verb constructions and prepositional verbs. 13. Complementation. 14. Beyond the clause.

Now let us look at the chapter headings in a very large grammar, with 40 chapters, the average length of the chapters being 33 pages.

Toqabaqita, Malaita, Solomon Is. (Lichtenberk 2008)

1. Introduction. 2. Phonology. 3 Grammatical profile. 4. The verb phrase. 5. The particles in the verb complex. 6. The noun phrase. 7. The noun-phrase internal particle group. 8. Possessive and associative noun phrases. 9. Nominalizations and deverbal nouns. 10. The prepositional phrase. 11. Coordination of noun phrases and prepositional phrases. 12. Compounding. 13. The demonstratives and the demonstrative adverbs. 14. Constructions with inclusory pronominals. 15. Tense and aspect. 16. The sequential subject markers. 17. Negation. 18. Mood. 19. Interrogatives. 20. Imperatives. 21. Low-individuation of participants derivations. 22. Reciprocal and related situation types. 23. Self-contained situations. 24. Unrestricted choice. 25. Comparisons of inequality. 26. Locational, existential, and possessive sentences. 27. Verbless sentences. 28. Coordination of clauses. 29. Complement clauses. 30. Relative clauses. 31. Conditional sentences. 32. Concessive clauses. 33. Reason clauses, purpose clauses, and purpose nominalizations. 34. Consequence clauses. 35. Temporal relations. 36. Deranked subordinate clauses. 37. Direct speech. 38. Topicalization. 39. Focusing. 40. Toqabaqita, Solomon Islands Pijin and English. Appendix: Texts. Notes. References. Index.

At first glance it may appear that the Toqabaqita grammar treats five times as many different topics as the PhD grammars containing just eight chapters. Certainly, the range of topics is considerably greater. However, the main difference is not so much the range of topics as the depth of coverage, the amount of detail to be found in the large grammar. Topics that occupy one section of shorter grammars, or that perhaps receive only passing mention, get a whole chapter to themselves in the Toqabaqita grammar.

With respect to the ordering of chapters Rice (2006) quotes a general principle from the Cambridge grammar series:

Basically, if an analytic decision concerning category X needs to refer some facts concerning category Y, then the chapter dealing with Y should be positioned before that dealing with X.

However, things are often not quite that simple. Whichever order is chosen, there is bound to be a need for extensive cross-referencing.

4.3 SHOULD A GRAMMAR BE CONCERNED WITH HOW TO SAY THINGS IDIOMATICALLY? I have always been concerned by the fact that knowing the core grammatical rules of a language will not allow you to speak that language idiomatically, saying things the way a native speaker says them. Writers of reference grammars traditionally are not much interested in what things people commonly say in a language and how they say these things, beyond the question of what is grammatical. Only a very few reference grammars contain sections on 'How to say things', ways of talking about everyday topics, even when there are very clear rules or conventions for doing so. Try looking in a grammar for an account of the conventions for telling the time of day. An exception in my sample of grammars of Pacific Island languages is Robert Bugenhagen's grammar of Mangap-Mbula, which has an appendix that begins:

It would be a tragedy to read through all the listings of word classes, rules and trees of a grammar and still not have any idea how to express some of the everyday things which make up a large part of people's normal conversation. (Bugenhagen 1995:374)

Under this heading the author goes on to treat 1. Existential and presentative constructions. 2. Location. 3. The weather, time of day, etc. 4. Movement. 5. Ownership and other types of relationships. 6. Emotion. 7. Conveyance. 8. Affectedness, different types of causation. 9. Speech/quotations. 10. Want, try, believe, promise, persuade, know, see and think (verbs that take sentential complements).

Bugenhagen wrote his grammar in the early 1990s. In today's grammars you are likely to find some of these topics treated in the main body of the grammar, because they concern constructions where morphosyntax craves correlation with function. Bugenhagen's topics 1, 2, 5, 8, 9 and 10 fall into this category. Others are less likely to be found in a grammar. It would be an interesting exercise to compare a sample of grammars and see how many of them give some prominence to ways of talking about particular subject matters, such as weather, or emotions, or conveyance.

5. THE FINER DETAILS OF PRESENTATION

5.1 SOME QUESTIONS. The previous section took a broad view of the contents of a grammar. Let us now consider some of the finer, nitty-gritty details of presentation. For example: Should the morphosyntactic part of the grammar begin with an overview of the main types of clause constructions before getting into the details of word formation and word classes? Should there be a separate chapter on word classes or should each word-class be introduced when dealing with the type of construction headed by that word-class (e.g. verbal clauses, noun phrases, etc)? Should each chapter begin with a summary of what it is about? What notational devices should be used, e.g. tree diagrams or square brackets to show constituent structure? How many examples should one give to illustrate a particular point? And for each construction type should one try to include a compact generative grammar in the form of explicit, concise formulae that will (aim to) predict all and only the possible strings, or should one be content to discuss constructions in a more informal way, with limited use of generative formulae?

5.2 SOME DOS AND DON'TS. A recent issue of *Studies in Language* (Payne and Webber 2006) contains papers by a number of linguists reflecting on issues in grammar writing. Michael Noonan's contribution (Noonan 2006) includes a list of dos and don'ts, based on suggestions provided by various experienced grammarians. The dos and don'ts concern, not so much the things that everyone agrees should be done in a grammar, like a thorough description of the morphology and syntax, but some of the things that tend to get left out or done not so well. Noonan divides them into three broad categories: User friendliness, descriptive adequacy and comprehensiveness. A summary of some of these prescriptions follows.⁶

⁶ I have omitted a few of Noonan's items, including those that we have already discussed, and have renumbered them accordingly.

- **User friendliness**

1. Provide a detailed index and table of contents. This helps readers find information.
2. The text should be divided into numbered and titled sections and subsections, with ample cross-referencing. Important terms should be highlighted by boldface.
3. Provide plenty of examples. Made up examples are appropriate for presenting information about basic constructions [when the writer is certain of their correctness: AP] but naturally occurring examples should otherwise be used.
4. Provide interlinear glosses (morpheme by morpheme) translations, as well as free translations for all examples.
5. A typological sketch of three to five pages should be included at the beginning of the grammar. This gives the reader a quick overview of the most important elements in the grammar.

- **Descriptive adequacy**

6. Use standard IPA characters to present information about the phonetics of the language.
7. Give instrumental documentation of the acoustic properties of vowels, the duration of segments, and tone and pitch accent systems.
8. Provide a full description of segmental and suprasegmental contrasts and the evidence for these.
9. Provide a full description of distributional patterns of elements of the phonology, in terms of syllables, words and whatever other units are relevant.
10. For morphologically complex languages, provide not just lists of affixes but tables with full paradigms showing combinations of all relevant morphemes.
11. Define grammatical categories used in the grammar. Don't assume that word classes and grammatical relations (subject, direct object etc) are givens.
12. The choice of labels for grammatical features is not as important as a thorough presentation of the facts.
13. It's better to admit ignorance of a poorly understood grammatical feature than to say nothing about it. Saying nothing can be misread as indicating that the feature is lacking.
14. The absence of a feature should be noted, if that feature might be expected to occur on areal, genetic or typological grounds.
15. Indications of frequency of grammatical elements and constructions should be provided where appropriate.
16. It is best to describe morphology mainly with a form-to-function orientation and syntax mainly with a function-to-form orientation.
17. A vocabulary of all the lexemes which occur in the grammar should be provided.
18. A collection of texts, at least 20-30 pages, with morpheme glosses and translations, should be included.

- **Comprehensiveness**

19. The writer should consult survey questionnaires and well-regarded grammars to make sure that important topics are not missed.
20. The grammar should contain information about genetic and areal affiliations of the language.

21. The grammar should contain information about how the data was obtained and about the sociolinguistic context. The latter should include the number and location of speakers, the age demographics of language use, the degree and nature of multilingualism, degree of literacy, etc.
22. There should be ample reference to previous scholarship on the language and the culture of the community.

Noonan ends with two items that are desiderata rather than requirements:

23. A good dictionary is a powerful adjunct to a good grammar. It will contain much grammatical information supplementing that in the grammar.
24. Where practical, audio and video recordings should be made of various genres.

Most grammarians would surely agree with most of these points although the question arises whether all of prescriptions 1-22 should apply to PhD grammars, which should not aim to cover everything in depth. The most contentious claim among these might be 16, saying that it is best to describe morphology mainly with a form-to-function orientation and syntax mainly with a function-to-form orientation. This is a useful general rule of thumb but the issues here are complex and one should be wary of hard and fast recommendations.

To Noonan's list I would add:

25. Begin each chapter with a summary of what it is about.
26. When describing the internal structure of complex constituents (e.g. noun phrase, verb, verbal complex, transitive clause) give explicit statements of combinatorial possibilities.
27. For the published version of the dissertation, include an index giving page references for key topics and terms.

6. HOW NOT TO COMPLETE A DISSERTATION: SOME BAD STRATEGIES IN DATA-GATHERING AND WRITING UP. A proportion of PhD students never finish their grammars. In my experience, the reason is seldom loss of interest, except when this is allied to long-term depression. Grammarians generally love their work. Sometimes the failure to finish is because of external factors — illness, lack of money, etc. — but leaving these factors aside, the main causes of incompletions in my experience are the following:

1. *Database addiction.* In many cases the symptoms of database addiction can also be spotted quite early. The student says he or she cannot begin serious writing until an extensive corpus of data has transcribed, annotated and analysed, or otherwise processed. At first the advisor accepts this line but time goes by and the database keeps growing and growing, and very little in the way of chapter drafts appears, the advisor realizes that something is wrong. The database has become an excuse for not writing up analyses, and has led to a form of writer's block.
2. *Theory addiction.* A commentator on a draft of this chapter refers to dissertation writers who, chronically insecure about their command of theory, fail to complete

because “there’s always one more paper to read, or the theory shifts and the writer feels compelled to revise”.

3. *Perfectionism*. Usually, it doesn’t take long to spot a perfectionist. He or she is someone who is reluctant to show you any chapter drafts until they are close to perfect and who consequently makes very slow progress. The perfectionist spends an inordinate amount of time worrying about analyses and revising drafts, or talking about the problems of analysis, without actually getting much onto paper.
4. *The all chapters at once strategy*. In this strategy, you write fragments of most or all chapters at once. (I use ‘chapter’ loosely for ‘section on a major topic’.) First you decide on a format for the grammar including chapter topics and subtopics. Then you search your database for examples that are relevant to each of the topics and sub-topics and insert these examples in the relevant subsections. Then you start to write bits of commentary on these examples. After a while you have written bits about sections 2 and 7 of chapter 3, sections 3 and 5 of chapter 4 and sections 2, 4 and 6 of chapter 6, and so on. No chapter is ever more than half finished. After a while the advisor discovers this is happening, tells you it is a bad method, and asks you to complete a draft of one chapter at a time so the advisor can give feedback. You say this is impossible because chapters 3, 4 and 6 are interdependent. Or you promise to change your ways but cannot actually bring yourself to do so.

A milder variant of type 3 is where the student submits drafts of a sequence of chapters to the advisor(s), say 2-3, and gets back comments on each but presses ahead with writing drafts of chapters 4, 5 and 6, without looking carefully at these comments. The upshot is that the faults of writing style, analysis, etc. exhibited in chapters 2-3 are repeated in 4-6.

Of course bad practices are not long-term problems if they can be changed. But when they reflect deeply ingrained character traits, it is a different story.

7. THE ROLE OF THE DISSERTATION ADVISOR(S). Up till now I have assumed that the voice of the advisor is constantly to be heard addressing the various issues discussed, but I have seldom made explicit reference to the duties of the dissertation advisor(s) (whom we call in the Antipodes, the supervisor(s)). Perhaps I should be more explicit. In the department at the Australian National University (ANU) where I have taught for the last 20 years, there is normally a panel of three supervisors. Two are primary supervisors: there is a chief supervisor or chair of the panel and a co-supervisor who has almost equal responsibilities with the chair. Both are expected to advise the student at all stages of the project. The third supervisor’s main duty is to read the completed draft, once the two main supervisors are reasonably satisfied with this. At the ANU the thesis is sent to a set of external examiners, usually three. The student’s thesis panel nominates the examiners but plays no part in the actual examination of the dissertation. In the USA the system is different. The committee consists of a chair, who is the principal advisor, and several other members, who play lesser roles but typically read the dissertation when it is complete. The completed dissertation is examined in-house by the student’s committee.

There are times where the advisor(s) and student need to be in regular contact to discuss issues and other times where the student can proceed independently. In the beginning, when the details of the project is being planned, there should be regular exchanges. The same applies in the later stages, when the student is writing the grammar.

The task of describing a little-known language normally requires extended fieldwork, often in a remote place. In the case of students doing fieldwork in a context of which they have no previous experience it is desirable, but not always possible, for an advisor to accompany the student into the field on the first trip and stay for a time.

Faced with a student who expresses a wish to do a grammar for his/her PhD, I generally begin by discussing the pros and cons of such a choice as a PhD topic and whether the student has the training for it. No one should try to write a grammar without a good grounding in the fundamentals of descriptive linguistics. Ideally, one should have taken courses in and read widely in all the basic fields of descriptive and typological linguistics, and looked carefully at number of reference grammars. One should read what literature there is on the target language, and on the language family to which it belongs, and should also read such literature as there is on the way of life of the communities speaking these languages.

The advisor must assess whether a student is suited by training, abilities, work habits and temperament to complete all the steps in a grammar-writing project. Obviously, one should not encourage or accept a student to do such a project if s/he seems unsuited to the task. However, in the absence of a track record of having previously completed a similar task, e.g. a master's thesis, this is always a bit of a gamble. There are various stages where students can lose their way.

Unsurprisingly, the most successful PhD grammar writers are those who have all the desirable qualities: they have sharp and enquiring minds, are well-trained in theory and skilful in data-collecting and analysis. They are enthusiastic and hard-working, enjoy fieldwork and are well-organized. They have the good judgment and flexibility to recognize and accept good advice and to question that which is dubious.

One remembers with a warm glow those students who had all these qualities. One young woman, in particular, was good at everything and dedicated to her language but where she really stood out was in the way she organized her two main advisors (myself and my colleague Malcolm Ross) during the writing up stages. One would have to say she took charge and, in the nicest way, controlled her advisors like puppets on a string. She lived 400 km west of Canberra but each month would come to the city and stay for a week near the university and make appointments to see each of us separately. She would email each advisor in advance with a set of very clear questions to be discussed at the meetings. If there was a conflict of advice between advisors this would lead to further set of well formulated questions. On each visit she would leave chapter drafts with each advisor and expect us to have our comments on these drafts ready for her next visit. The two advisors never dealt with the same draft at the same time. The student arranged things so that advisor A read the first draft and, after she had revised according to A's comments, advisor B got the revised draft. All her requests were made with such politeness, and every piece of advice was used so well that we did not mind being exploited with maximum efficiency. It came as no surprise that an excellent dissertation was completed in quick time and published soon

after, or that in the many years since her PhD was completed she has maintained a close and productive association with the community whose language she described.

REFERENCES

- Aikhenvald, Alexandra A. 2008. *The Munambu language of East Sepik, Papua New Guinea*. Oxford: Oxford University Press.
- Ameka, Felix, Alan Dench and Nicholas Evans, eds. 2006. *Catching language: the standing challenge of grammar-writing*. Berlin: Mouton de Gruyter.
- Bowden, John. 2001. *Taba: Description of a South Halmahera language*. Canberra: Pacific Linguistics.
- Bowern, Claire. 2008. *Linguistic fieldwork: a practical guide*. New York: Palgrave MacMillan.
- Bugenhagen, Robert. 1995. *A grammar of Mangap-Mbula, an Austronesian language of Papua New Guinea*. Canberra: Pacific Linguistics.
- Dixon, Robert M.W. 2009-10. *Basic linguistic theory*. vol. 1. *Methodology*. (2009); vol. 2. *Grammatical topics* (2011); vol. 3. *Further grammatical topics*. (2012). Oxford/New York: Oxford University Press.
- Dol, Philomina. 2007. *A grammar of Maybrat, a language of the Bird's Head Peninsula, Papua Province, Indonesia*. Canberra: Pacific Linguistics.
- Dryer, Matthew S. 2006. *Descriptive theories, explanatory theories, and basic linguistic theory*. In Ameka et al., pp. 207-234.
- Eades, Domenyk. 2005. *A grammar of Gayo, a language of Aceh, Sumatra*. Canberra: Pacific Linguistics.
- Ezard, Bryan. 1997. *An Austronesian language of the Milne Bay area, Papua New Guinea*. Canberra: Pacific Linguistics.
- Farr, Cynthia. 1999. *The interface between syntax and discourse in Korafe, a Papuan language of Papua New Guinea*. Canberra: Pacific Linguistics.
- François, Alexandre. 2001. *Contraintes de structures et liberté dans l'organisation du discours. Une description du mwotlap, langue océanienne du vanuatu*. Thèse doctoral. Université Paris IV–Sorbonne.
- Huddleston, Rodney and Geoffrey Pullum. 2002. *The Cambridge grammar of the English Language*. Cambridge: Cambridge University Press.
- Hyslop, Catriona, 2001. *The Lolovoli dialect of the North-East Ambae language*. Canberra: Pacific Linguistics.
- Jones, Alan. 1998. *Towards a lexicogrammar of Mekeo (an Austronesian language of western central Papua)*. Canberra: Pacific Linguistics.
- Klamer, Marian. 1994. *Kamera, a language of Eastern Indonesia*. Canberra: Pacific Linguistics.
- Lichtenberk, Frantisek. 2008. *A grammar of Toqabaqita*. (2 vols) Berlin: Mouton de Gruyter.
- Mosel, Ulrike. 2011. Morphosyntactic analysis in the field: a guide to the guides. In Nicholas Thieberger, ed. *The Oxford handbook of linguistic fieldwork*, 72-89. Oxford: Oxford University Press.

- Mosel, Ulrike and Even Hovdhaugen. 1992. *Samoan reference grammar*. Oslo: Scandinavian University Press.
- Noonan, Michael. 2006. *Grammar writing for a grammar reading audience*. In Payne and Weber, 351-365.
- Obata, Kazuko. 2003. *A grammar of Bilua, a Papuan language of the Solomon Islands*. Canberra: Pacific Linguistics.
- Payne, Thomas E. 1997. *Describing morphosyntax. A guide for field linguists*. Cambridge: Cambridge University Press.
- Payne, Thomas E and David J Weber, eds. 2006. Perspectives in grammar writing. Special issue of *Studies in Language* 30(2). Benjamins: Amsterdam/Philadelphia.
- Pensalifini, Rob. 2003. *A grammar of Jingulu, an aboriginal language of the Northern Territory*. Canberra: Pacific Linguistics.
- Quick, Phil. 2007. *A grammar of the Pendau language of central Sulawesi, Indonesia*. Canberra: Pacific Linguistics.
- Quirk, Randolph, Sidney Greenbaum, Jeffrey Leech and Jan Svartvik. 1974. *A grammar of contemporary English*. London: Longmans.
- Rice, Keren. 1989. *A grammar of Slave*. Berlin: Mouton de Gruyter.
- Rice, Keren. 2006. *A typology of good grammars*. In Payne and Weber, 385-415.
- Sharp, Janet. 2004. *Nyangumarta, a language of the Pilbara region of Western Australia*. Canberra: Pacific Linguistics.
- Teng, Stacey Fang-ching. 2008. *A reference grammar of Puyuma, an Austronesian language of Taiwan*. Canberra: Pacific Linguistics.
- van den Heuvel, Wilco, 2006. *Biak. Description of an Austronesian language of Papua*. Utrecht: LOT.
- van Klinken, Catharina. 1999. *A grammar of the Fehan dialect of Tetun*. Canberra: Pacific Linguistics.
- van Staden, Miriam. 2000. *Tidore, a linguistic description of a language of the North Moluccas*. PhD dissertation, Universiteit Leiden.
- Wegener, Claudia, 2008. *A grammar of Savosavo, a Papuan language of the Solomon Is.* MPI Series in Psycholinguistics 51. Niemegen: Max Planck Institute for Psycholinguistics.

Andrew.Pawley@anu.edu.au