

How to write a grammar of an undescribed language: introductory issues

Nick Evans
Nicholas.evans@anu.edu.au

1. Introduction

2. The goal: what would a perfect grammar be like?

Some ideals (not all mutually compatible)

The Paninian ideals:

- explicitness (of rules and terms)
- brevity
- architectural elegance

The structuralist's ideals:

- respecting the 'distinctive genius' of the language
- treatment of the grammar as a system - i.e. examining consequences of interactions between rules, categories
- need to give language-internal justification for categories, labels etc.
- need for grammatical description to be embedded in, and mutually consistent with, the Boasian trilogy of grammar + texts + dictionary

The generativist's ideals:

- recognition of the recursive nature of language, of the need to account for an infinitely large corpus by generative means
- need for formal precision of rules, so as to examine their interaction (essentially this continues the Paninian ideals)

The typologist's ideals:

- typological consistency of terminology, i.e. need to anchor definitions against cross-linguistically accepted usage –
(when does an adjective stop being an adjective? what counts as aspect?
when should one stop using a common phonetic symbol and adopt a more exotic one?)
- interrogability on any conceivable typological topic e.g. aspect, interrogative and indefinite pronouns, reciprocals – paired with close attention to semantic detail
Favours function-to-form (semasiological) organisation over form-to-function

(onomasiological) organisation – see Mosel (2006), Cristofaro (2006), Gabelentz (1891)

- Virtually no grammars do this (except Leech & Svartvik 1975's 'Communicative grammar of English')
- Some attempts to get around this are
- the encyclopaedic organisation in Newman's Hausa grammar,
 - clever use of indexing
(e.g. Haspelmath's grammar of Lezgian)
 - use of hypertext to allow multiple routes of entry
(developing technology; I don't know of a good example yet, but see Zaefferer 2006)
 - inconsistent organisation (e.g. Evans 1995) that moves back and forward between formal and functional patterns; often grammars adopt a more function-based organisation once they get to complex sentences

The documentarist's ideals:

- need to anchor description in a publically-available, verifiable corpus (cf Heath 1984, Thieberger 2007)
- associated need to separate tasks of documentary and descriptive linguistics (Himmelmann 1998, Gippert, Himmelmann & Mosel 2006)
- accountability of grammar to a whole, balanced, multi-genre corpus
- need to represent rationale for translations, distinguishing utterance-translations from other possible translations (see Evans & Sasse 2007 for some discussion)

The sociolinguist's ideals:

- need to represent sociolinguistic variation – analytic choice as to whether to write a description of a single system (e.g. one dialect) or of a diasystem (see Evans 2003 for an attempt)
- an unresolved problem is how to write a grammar of a whole diasystem

The linguistic anthropologist's ideals:

- need to portray language in its cultural context (see Hill 2006)

The pedagogical / communicative ideal:

- need to write the grammar in a way that it can be read!
- organisation: pedagogical progression vs. reference-grammar progression

rules: compressed / symbolised / theory-specific formulation vs. consensus, informal formulation

audience / presupposition: how much can be taken for granted in terms of glossing, definition of analytic concepts? (see Bauer et al 1997 for an example that tries to explain all concepts in layman's terms before progressing to the main grammar)

3. The process: moving closer to the mirage

3.1 Helical process in 2 senses

I.e. need to work on

- texts
- dictionary
- grammar

at the same time

But also, within the grammar, you need to work on

- phonetics
- phonology
- morphology
- syntax
- semantics of categories
- system of word classes etc.

and each of these is dependent on the others

This is the most intellectually demanding part of writing a grammar: the need to work on hundreds of different problems, in parallel mode, and keep track of your analytical decisions about each of them

Practicalities

– need for long stretches of uninterrupted concentration time

– need to keep track of provisional analyses at all stages

(they won't always be internally consistent)

therefore important to have a working structure e.g. a provisional table of contents

so you know where to put things – I find that analysing a single sentence or entering a single dictionary item can make me want to jot down thoughts sometimes in five or six different parts of the emerging grammar

– the more the grammar emerges, the more important it becomes to keep going over texts to check that everything is accounted for. So part of my time management in the later phases is always to make some time to go over a new text and see what new things come up

- debate on role of software tools like Shoebox/Toolbox
 - in favour
 - enforces consistent analysis,
 - allows quick checking of whole corpus,
 - interface with dictionary-building and text-glossing software
- against
- favour premature closure (premature hardening of analysis)
 - take the analysis out of your head into a computer
 - ultimately a good grammar needs to be entirely in your head

3.2 Isolated vs. nodal problems

Nodal problems – those where there is a dense interaction with many rules / phenomena, so that there are complex interdependencies between analyses in different parts of the grammar

Examples:

Word classes (in any language)

E.g. adopting word-class distinctions in one part of the grammar then affects

- applicability of rules over items
- whether treatments of conversion are needed

Nature of inflection vs. derivation

In Kayardild problems were:

- (a) multiple inflection
(so final position couldn't be used as a criterion for case)
- (b) existence of word-class changing inflections, so that

Nature of core grammatical organisation (is there a subject? or a topic-based system? or both? this impacts on analysis of voice, relativisation, agreement and many other things)

The more such nodal problems there are, the harder it is to reach an analysis – see Himmelmann (2006) on the lack of an agreed analysis of Tagalog (and Philippine languages more generally) after several

more than 250 years of more or less continued grammatical analysis do not seem to have been long enough to establish a widely accepted basic grammaticographic practice for Tagalog (or any other Philippine language, for that matter). (Himmelmann 2006:487)

In practical terms, you need to develop a feeling for which problems are relatively isolated – and can be worked on one at a time, and build up some bulk to your description – and which are more nodal, and will take longer (and much greater mental efforts) to solve

3.3 The importance of examples

You can't invent everything from scratch. The role of good examples should be to give you good models for some of what you need to do, leaving you free to devote your creative originality to the parts of the language which don't fit existing descriptions

Reading – the seven pillars

- (a) grammars of your mother tongue (in your own and other languages)
- (b) grammars of other languages you know well (again, written in a variety of languages)
- (c) grammars of 'exotic' languages which you may or may not have studied, across a variety of linguistic types

Among my favourites:

Haspelmath – Lezgian
Mosel & Hovdhaugen – Samoan
Kruspe – Semelai
Lichtenberk – Manam
Newman – Hausa
Osumi – Tinrin
Suttles – Musqueam
Tamura – Ainu
Valentine – Nishnaabemwin

as well as the Australian grammars mentioned under (d)

- (d) grammars of languages related to the one you are describing
E.g. in writing my grammar of Kayardild, I was heavily influenced by Dixon's 2 great Australian grammars (Dyirbal, Yidiny); I had got to know the first one while taking a course on Dyirbal with him, but actually the Yidiny grammar is much richer and deeper
Also important to me were:
 - Austin (1981) on Diyari, for its conciseness and precision

- Donaldson (1980) of Ngiyambaa for its semantic and cultural sensitivity, and its way of presenting examples in context
- shorter (100-200 pp.) grammars in the *Handbook of Australian Languages* series, particularly Keen's grammar of the closely related language Yukulta
- Hale's short but brilliant grammatical sketch of Lardil (only around 50 pp.) which was important in dealing with 2 key problems that also appear in Kayardild (tense-sensitive case endings, and verbal case) (finally appeared as Hale 1997)
- Hale's (1982) article on Warlpiri main clauses

(e) 'essays' that set out the architectural principles of how particular grammars are organised – without necessarily attempting a complete analysis (e.g. Launey 1994 on the 'omnipredicative' organisation of Nahuatl was very important to me when I was thinking about how to organise my grammar of Bininj Gun-wok); another good example of this genre is Aissen (1987) on argument structure in Nahuatl

(f) typological treatments of particular topics, e.g. Haspelmath's (1997) book on Indefinite Pronouns, or the treatment of particular topics (negation, relative clauses, Tense/Aspect/Mood, Deixis in Shopen's 3-volume collection (1985, 2007)

(g) any papers on particular problems in your language or languages closely related to it

3.4

For many other good tips on how to incorporate the challenge of grammar-writing into your life see Weber (2005)

References

- Ameka, Felix, Alan Dench & Nicholas Evans (eds.) *Catching language: the standing challenge of grammar-writing*. Berlin: Mouton de Gruyter.
- Austin, Peter. 1981. *A grammar of Diyari*. Cambridge: CUP.
- Bauer, Winifred (with William Parker and Te Kareongawai Ewans). 1997. *The Reed reference grammar of Maori*. Auckland: Reed.
- Cristofaro, Sonia. 2006. The organization of reference grammars: A typologist user's point of view. In Felix Ameka, Alan Dench & Nicholas Evans (eds.) *Catching language: the standing challenge of grammar-writing*. Berlin: Mouton de Gruyter. Pp. 137-170.
- Dixon, R. M. W. 1972. *The Dyirbal Language of North Queensland*. Cambridge: CUP.
- Dixon, R.M.W. 1977. *A Grammar of Yidj*. Cambridge: CUP.

- Donaldson, Tamsin. 1980. *Ngiyambaa, the language of the Wangaaybuwan*. Cambridge: CUP.
- Evans Nicholas. 1995. *A Grammar of Kayardild*. Berlin: Mouton de Gruyter.
- Evans, Nicholas. 2003. *Bininj Gun-wok: a pan-dialectal grammar of Mayali, Kunwinjku and Kune*. (2 volumes). Canberra: Pacific Linguistics.
- Evans, Nicholas & Alan Dench. Introduction: catching language. In Felix Ameka, Alan Dench & Nicholas Evans (eds.) *Catching language: the standing challenge of grammar-writing*. Berlin: Mouton de Gruyter. Pp. 1-39.
- Evans, Nicholas and Hans-Jürgen Sasse. 2007. [On-line reprint, slightly revised, of Evans & Sasse 2004]. Searching for meaning in the library of Babel: field semantics and problems of digital archiving. *Archives and Social Studies: A Journal of Interdisciplinary Research* Vol. 1.0:260-320. On-line publication available at http://socialstudies.cartagena.es/index.php?option=com_content&task=view&id=38&Itemid=33
- Gabelentz, Georg von der. 1891. *Die Sprachwissenschaft. Ihre Aufgaben, Methoden und bisherigen Ergebnisse*. Leipzig.
- Gippert, Jost, Nikolaus P. Himmelmann & Ulrike Mosel (eds.) 2006. *Essentials of Language Documentation*. Berlin: Mouton de Gruyter.
- Hale, Ken. 1982. Some essential features of Warlpiri verbal clauses. In S.M. Swartz (ed.) *Papers in Warlpiri grammar in Honour of Lothar Jagst*. Darwin: Summer Institute of Linguistics. Pp. 217-235.
- Hale, Ken. 1997. Remarks on Lardil phonology and morphophology. In Ngakulmungan Kangka Leman (compilers), *Lardil dictionary*. Mornington Shire Council, pp. 12-56.
- Haspelmath, Martin. 1993. *A grammar of Lezgian*. Berlin: Mouton de Gruyter.
- Heath, Jeffrey. 1984. *A functional grammar of Nunggubuyu*. Canberra; AIAS.
- Hill, Jane H. 2006. Writing culture in grammar in the Americanist tradition. In Ameka, Dench & Evans (eds.), pp. 609-628.
- Himmelmann, Nikolaus. 2006. How to miss a paradigm or two: multifunctional ma- in Tagalog. In Ameka, Dench & Evans, pp. 487-527.
- Kruspe, Nicole. 2003. *A grammar of Semelai*. Cambridge: CUP.
- Launey, Michel. 1994. *Une grammaire omniprédicative*. Paris: CNRS Editions.
- Leech, Geoffrey and Jan Svartvik. 1975. *A Communicative Grammar of English*. London: Longman.
- Lichtenberk, Frantisek. 1983. *A grammar of Manam*. Honolulu: University of Hawaii Press.
- Mosel, Ulrike. 2006. Grammaticography: the art and craft of writing grammars. In Ameka, Dench & Evans (eds.), pp. 41-68.
- Mosel, Ulrike & Einar Hovdhaugen. 1992. *Samoan Reference Grammar*. Oslo: Scandinavian University Press.
- Newman, Paul. 2000. *An encyclopaedic grammar of Hausa*. Yale: Yale University Press.
- Osada, Toshiki. 1992. *A Reference Grammar of Mundari*. Tokyo: Institute for the Languages and Cultures of Asia and Africa.
- Rice, Keren. 2005. A typology of good grammars. *Studies in Language* 30.2:385-415.

- Shopen, Timothy. 1985 (rev. edn. 2007). *Language Typology and Syntactic Description*. (3 vols)/ Cambridge: Cambridge University Press.
- Suttles, Wayne. 2004. *Musqueam Reference Grammar*. Vancouver: UBC Press.
- Tamura, Suzuko. 2000. *The Ainu language*. Tokyo: Sanseido.
- Thieberger, Nicholas. 2007. *A grammar of South Efate*. Honolulu: University of Hawaii Press.
- Valentine, Randolph. 2001. *Nishnaabemwin Reference Grammar*. Toronto: University of Toronto Press.
- Weber, David J. 2005. Thoughts on growing a grammar. *Studies in Language* 30.2:417-44.
- Zaefferer, Dietmar. 2006. Realizing Humboldt's dream: Cross-linguistic grammatography as data-base creation. In Felix Ameka, Alan Dench & Nicholas Evans (eds.) *Catching language: the standing challenge of grammar-writing*. Berlin: Mouton de Gruyter Pp. 113-136.