Trends in Linguistics
Studies and Monographs 167

Editors

Walter Bisang

(main editor for this volume)

Hans Henrich Hock

Werner Winter

Mouton de Gruyter
Berlin · New York

Catching Language

The Standing Challenge of Grammar Writing

edited by

Felix K. Ameka

Alan Dench

Nicholas Evans

Mouton de Gruyter
Berlin · New York

Mouton de Gruyter (formerly Mouton, The Hague)
is a Division of Walter de Gruyter GmbH & Co. KG, Berlin.

© Printed on acid-free paper which falls within the guidelines
of the ANSI to ensure permanence and durability.

Library of Congress Cataloging-in-Publication Data

Catching language : the standing challenge of grammar writing / edited
by Felix K. Ameka, Alan Dench, Nicholas Evans.
p. cm. — (Trends in linguistics. Studies and monographs ; 167)
Includes bibliographical references and indexes.
ISBN-13: 978-3-11-018603-1 (alk. paper)
ISBN-10: 3-11-018603-9 (alk. paper)
I. Language and languages — Grammars — Authorship. I. Ameka,
Felix K. II. Dench, Alan Charles. III. Evans, Nicholas, 1956–
P53.412.C38 2006
808'.066418—dc22

2006022868

ISBN-13: 978-3-11-018603-1
ISBN-10: 3-11-018603-9
ISSN 1861-4302

Bibliographic information published by the Deutsche Nationalbibliothek

The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie;
detailed bibliographic data are available in the Internet at <http://dnb.d-nb.de>.

© Copyright 2006 by Walter de Gruyter GmbH & Co. KG, D-10785 Berlin
All rights reserved, including those of translation into foreign languages. No part of this
book may be reproduced or transmitted in any form or by any means, electronic or mechan-
ical, including photocopy, recording or any information storage and retrieval system, with-
out permission in writing from the publisher.
Cover design: Christopher Schneider, Berlin.
Printed in Germany.

Preface

Despite a recent surge in awareness of the need to document little-known languages, we know of no other book that explores the manifold issues that face the author of a descriptive grammar. Courses in grammar-writing – as opposed to courses in the analysis of individual domains like phonology, morphology or syntax – are rarely offered, so that most grammar-writers need to develop their craft from scratch. We hope that the contributions in this volume, taken together, will help anyone intending to write a descriptive grammar to clarify their goals and methods along the path to producing the succinct, rigorous and sensitive masterpiece that each of the world's languages deserves. Because the composing of any complex work needs to approach the problem from two angles, that of the writer and that of the reader, we have made sure that both points of view are represented here: most contributors have written descriptive grammars themselves, but some represent grammar users rather than grammar writers. Many, of course, regularly move between the two roles.

We are grateful to the contributors who did not wane in their commitment despite the changing circumstances surrounding the project. We are indebted to Luisa Miceli and Hywel Stoakes, without whom the production could not have been achieved, and to Alice Gaby and Eva Fenwick for compiling the index. The team at Mouton, especially Anke Beck and Birgit Sievert, deserve thanks not only for steering the book to its final production but also for their enduring confidence in the feasibility of the project.

The three of us were fortunate enough to have studied at The Australian National University at a time when it was one of the few places in the world where the norm for a doctoral dissertation was a grammatical description of an undescribed or little described language – a time when what has come to be known as “The ANU School of Grammar Writing” was fashioned. The ANU department was a melting pot of diverse personalities, ideas, methods and practices, all joined to the quest for understanding, recording and analyzing the forms, functions and meanings of a diverse range of languages. We dedicate this volume to all our teachers in this craft.

Contents

Preface	v
Introduction: Catching language <i>Nicholas Evans and Alan Dench</i>	1
Grammaticography: The art and craft of writing grammars <i>Ulrike Mosel</i>	41
Real descriptions: Reflections on native speaker and non-native speaker descriptions of a language <i>Felix K. Ameka</i>	69
Realizing Humboldt's dream: Cross-linguistic grammatography as data-base creation <i>Dietmar Zaefferer</i>	113
The organization of reference grammars: A typologist user's point of view <i>Sonia Cristofaro</i>	137
Calculus of possibilities as a technique in linguistic typology <i>Igor Mel'čuk</i>	171
Descriptive theories, explanatory theories, and Basic Linguistic Theory <i>Matthew S. Dryer</i>	207
Let the language tell its story? The role of linguistic theory in writing grammars <i>Keren Rice</i>	235
On describing word order <i>Randy LaPolla and Dory Poa</i>	269
Heterosemy and the grammar-lexicon trade-off <i>N. J. Enfield</i>	297

Field semantics and grammar-writing: Stimuli-based techniques and the study of locative verbs <i>Birgit Hellwig</i>	321
Taking a closer look at function verbs: Lexicon, grammar, or both? <i>Eva Schultze-Berndt</i>	359
Converbs in an African perspective <i>Azeb Amha and Gerrit J. Dimmendaal</i>	393
From Eurocentrism to Sinocentrism: The case of disposal constructions in Sinitic languages <i>Hilary Chappell</i>	441
How to miss a paradigm or two: Multifunctional <i>ma-</i> in Tagalog <i>Nikolaus P. Himmelmann</i>	487
The interplay of synchronic and diachronic discovery in Siouan grammar-writing <i>Robert L. Rankin</i>	527
The historical and cultural dimensions in grammar formation: The case of Modern Greek <i>Brian D. Joseph</i>	549
Polylectal grammar and Royal Thai <i>Tony Diller</i>	565
Writing culture in grammar in the Americanist tradition <i>Jane H. Hill</i>	609
List of contributors	629
Author index	632
Language index	643
Subject index	650

Introduction: Catching language

Nicholas Evans and Alan Dench

1. The challenge of grammar-writing

Descriptive grammars are our main vehicle for representing the linguistic structures of the world's 6,000 languages in all their bewildering variety. Each grammar seeks to bring together, in one place, a coherent treatment of how the whole language works, and therefore forms the primary source of information on a given language, consulted by a wide range of users: areal specialists, typologists, formal linguists, historical linguists, and members of the speech communities concerned. The writing of a descriptive grammar is a major intellectual and creative challenge, often taking decades to complete. It calls on the grammarian to balance a respect for the distinctive genius of the language with an awareness of how other languages work, to combine rigour with readability, to depict elegant structural regularities while respecting a corpus of real and sometimes messy material, and to represent the native speaker's competence while recognising the patterns of variation inherent in any speech community.

What does it mean to write a descriptive grammar? The grammarian seeks to capture the essential structural features of a language, and in codifying these presents them to a diverse and critical audience. That audience is primarily the community of linguists, itself quite broad, but it very often also includes language teachers, language learners and native speakers of the language. The grammar writer must be sensitive to this diversity. While letting the language 'tell its own story', and while letting its speakers 'speak for themselves, creating a record of spontaneous speech in natural communicative settings' (Mithun 2001: 53), the grammarian must still locate their description within the broad comparative concerns of linguistic typology and the received traditions of description within a language family. But at the same time they must remain open to new analyses that are either more comprehensive or insightful or less bound by a particular traditional descriptive template. A grammar should also be written with a re-

spect for the constantly evolving questions and concerns of formal linguistic modelling – both in response to the new discoveries about the nature of grammatical phenomena revealed by this enterprise and as a potential proving ground for competing models of human language. Inevitably, given the number of constraints that must be balanced, there is no single ideal solution to this challenge, and lucky is the language where it is possible to consult several first-rate grammars bringing different interests and analytic traditions to bear on the same language.

In catching the elegance and intricacy of the language they are describing, the grammarian must incessantly struggle with what should be in the grammar and what should be left out, of where the boundaries lie between grammar and lexicon, between linguistic description and ethnography, between one linguistic variety and another, and between the current state of the language and its evolving history. The question of when to formulate explanation, and when to stop at description, always lurks close at hand. The boundary between description and prescription can also become blurred, since as soon as one admits speaker acceptability judgments as sources of data alongside a naturalistic corpus these may call forth prescriptive biases. Thus the descriptive grammar becomes a description of speakers' prescriptions as well as of their practice. This is of course particularly complicated for languages like Greek with a long and luminous literary tradition, and it can also be more difficult for native-speaker grammarians to dissociate description from their own prescriptive biases.

These questions appear and reappear, in different forms and applied to different case studies, throughout the book. In this introduction we summarize the issues in more detail referring briefly, as they come up, to relevant chapters with the author's name in brackets.

2. Language description, linguistic typology, and formal modelling

In our view there is a triadic and mutually complementary relationship between *descriptive linguistics* (of which the writing of grammars is but one part), *linguistic typology*, and *formal linguistics* – the subfield of linguistics concerned with developing precise, formalized models of syntax, phonology, semantics and other modules of the grammar. In 2.1. we discuss the relations between these three endeavours, while in 2.2. we show how this integrated view can avoid important conceptual confusions that are unfor-

tunately endemic in the field, particularly regarding the nature of explanation in linguistics.

2.1. Roles and interrelations

2.1.1. Descriptive linguistics

The job of descriptive linguistics is to describe individual languages as perceptively and rigorously as possible, with maximal accountability to a naturalistic corpus of data ideally collected within a broad program of language documentation (Himmelman 1998) to ensure that the full spectrum of language structures are represented. A naturalistic corpus is typically supplemented by speaker acceptability judgments that can help identify structures that are rare in natural speech, but are nonetheless within the competence of the native speaker (Rice), and can also weed out performance errors from the corpus. Particular 'semi-controlled' elicitation techniques such as stimulus sets may be used to increase the frequency of structures that might otherwise be too rare to be analysed with any real confidence (Hellwig). The chances of representativeness are also increased if the grammarian is a native speaker (Ameka), or takes the time and trouble to become fluent in the target language (Hale 2001, Mosel): the need to communicate on all topics in life is a potent reminder of which aspects of the language the linguist does not yet understand.

An emphasis on systemic integration and on recognizing the overall architecture of a language has dominated linguistics since the structuralist era. The idea of an essential 'genius of a language' drives descriptive grammarians to look for a characteristic overall *Bauplan* that makes sense of the language's particularities in an integrated way.¹ At the same time, grammarians need to keep an open mind on how far traits are in fact free to vary independently, or, following the structuralist dictum, hold each other together:

linguistic features that are easily thinkable apart from each other, that seem to have no necessary connection in theory, have nevertheless a tendency to cluster or to follow together in the wake of some deep controlling impulse to form (Sapir 1921/1949: 144)

Picking up on such architectonic principles may lead the grammarian to generalize the notions of systematic, paradigm-like organization to new

areas of grammar. Himmelmann's chapter adopts such an approach to the organization of the dimensions of voice, aspect/mood, and potentive vs stative category in Tagalog arguing that this is necessary to detect the poly-functionality of the prefix *ma-*.

Sensitive descriptions frequently discover quite new phenomena, previously unknown to general linguistics, which feed into the comparative enterprise of linguistic typology and may alert other descriptive linguists to their presence in other languages. A grammarian's appetite for grasping the unknown can be whetted by typological approaches whose 'calculus of possibilities' make explicit what phenomena are unknown to general linguistics (Mel'čuk). At the same time, relatively commonplace notions from general linguistics – e.g. 'subject', 'adjective', 'inflection' etc. – may require revision as their applicability to particular language descriptions is assessed, this feeding back into more careful formulation of their cross-linguistic definitions (Mel'čuk). The typology of word order is a particularly fraught example (LaPolla): typologists are keen to excerpt generalized statements about the principles of constituent order in particular languages, but such gross characterizations as 'SVO' may be completely at odds with the real language-specific principles that are active in, say, Chinese.

Descriptive grammars furnish accessible and well-defined accounts of a language that provide material for formal linguistics to model – a particularly important enterprise when complex interactions across different modules of the grammar are involved. At the same time, grammarians are challenged by formal approaches that raise theoretical stakes by claiming that certain phenomena are ruled out by universal constraints on human language. Autobiographically, Evans can confess that he was pushed to investigate certain aspects of polysynthesis in Bininj Gun-wok (particularly noun incorporation, and the semantics of pronominal prefixes) as a reaction to the claims about deep-level structural parallels between polysynthetic languages and English put forward in Baker (1988, 1995), resulting in a much more detailed description than he would have produced without this goad.

2.1.2. Linguistic typology

Typology is the subfield of linguistics concerned with developing a body of analytically compatible concepts or general conceptual framework valid across all the world's languages (Mel'čuk), through an unending cycle of

induction from individual languages (Croft 2001, Dryer) and deduction from general principles (Mel'čuk). It must integrate the vast library of grammatical descriptions (Zaefferer, Cristofaro) in a way that renders them at least broadly compatible. The difficulties in adjudicating terms across descriptions of different languages are nicely summarized in Dryer's chapter:

Whenever we employ distinct terms for phenomena in a language, we run the risk of obscuring the similarities with other languages. And whenever we employ familiar terms, we run the risk of obscuring difference.

Describing each language entirely on its own terms is a noble and galvanizing task, but unless grammarians orient their findings to what typologists know about the world's other languages, their grammars can all too easily become obscure, crabbed and solipsistic – such as Whorf's 'brilliant failure' in attempting to apply this principle to Hopi in an extreme and typologically naive form (Hill) – or at best half-veiled in the idiosyncrasies of specific areal or language-family-specific traditions (Amha and Dimmendaal, Chappell).

Typology constantly throws out new questions for descriptive linguists to investigate (Cristofaro) as it identifies parallels and contrasts in how different languages organize newly-discovered phenomena. As an example, see the discussion of complex verbs in Amha and Dimmendaal's chapter, which discuss comparable phenomena that have been named and treated rather differently by distinct descriptive traditions.

Typological approaches to the search for language universals, making use of methodologically rigorous sampling techniques, provide an important empirical base underpinning the search for universals of human language – an enterprise more commonly identified with formal linguistics and the Chomskyan paradigm. Ultimately, formal models of 'universal grammar' must be consistent with, and should elegantly account for, the typological diversity evidenced in the wealth of the world's languages. Conversely, typologists need to be sensitive to new questions that formal models raise about the properties of linguistic phenomena, and to lead these new questions back into their systematic confrontation with linguistic diversity.

2.1.3. *Formal linguistics*

Formal modelling seeks a more rigorous formulation of linguistic patterning than is generally possible in a normal descriptive grammar. The rigorous model-building that formal systems allow makes it possible to test exhaustively the interactions between the many different rules, of phonology, morphology, syntax, etc., that might be included in the average grammatical description. Descriptive linguists sensitive to this possibility have greatly benefitted from the use of formal theories as tools of discovery. In addition, the formal approaches associated with the generative tradition, in particular, have constantly drawn attention to new phenomena in the last four decades: anaphors, island constraints, different control properties, secondary predicates, WH-movement (Rice) and so on. That so many of these have been discovered relatively recently, in such a well-researched language as English, demonstrates the particular value of detailed formal modelling in showing that descriptive grammarians typically know far less about a particular language than they may have thought.

New representational devices, like phrase structure or metrical trees, or autosegmental representations, have played an important role both in the discovery of new phenomena and in simplifying their treatment in descriptive grammars (Rice). In this way, some elements of formal representation may assist the exposition in a descriptive grammar.

However, experience shows us that, sadly, there is generally an inverse relationship between the adoption in grammars of specific formalisms and their readability by linguists of different schools and at different times. The most enduring and accessible descriptions turn out to be those that employ natural language (rather than a formal representational system) as their descriptive metalanguage. The same inverse relationship applies to parts of descriptions, in the case of grammars that mix informal and formal sections (Dixon 1972, 1977)² or volumes (Kibrik 1977a,b).³ In each of these cases, the informal sections or volumes have stood the test of readability over time better than the formal ones.

The most readable grammars use an informal descriptive metalanguage that draws on the accumulated conceptual system known variously as 'general comparative grammar' (Lehmann 1989) or 'basic linguistic theory' (Dixon 1997, Dryer), whose cross-linguistic comparability is constantly being assessed and updated by linguistic typology and by formal linguistics. As Dryer's chapter cautions, formalism should not be confused with precision: 'the use of formalism does not guarantee precision if the formal-

ism is not defined', and a lack of formalism is not incompatible with precision provided that descriptive terms are given rigorous definitions and rules are formulated carefully.

It is helpful, in seeing how incremental descriptive traditions lead to ever more accurate descriptions of a given language, to distinguish three elements that interact to promote maximal falsifiability: (a) 'formal falsifiability', concerned with assessing the internal consistency of the model of the grammar; (b) 'empirical falsifiability', concerned with testing the adequacy of the model in accounting for the empirical facts of the language, and (c) 'accessibility of formulation', concerned with ensuring that the description is as accessible to as many scholars of the language as possible. A description that fails to employ a terminology comprehensible across schools and times risks being all dressed up with nowhere to go. However falsifiable it might be on mathematical/logical grounds, the audience of comprehending readers is too small to subject it to proper scrutiny against available data on the language. Thus it may never be adequately tested empirically: its inaccessibility renders it 'sociologically unfalsifiable', sidelining it from wider empirical testing. The grammar writer must thus balance the need to address questions arising from current formal investigations or from linguistic typology with the need for the description, written for a wider audience and for posterity, to transcend these concerns.

2.2. *Theory, description and explanation*

There is some tendency within linguistics to associate the notion of 'theory' with formal models exclusively. However, we argue that the notion of 'theory' is associated with all three parts of the triad described above. A descriptive grammar of a language is a theory of what it is possible to say, with specified meaning(s), in that language (Dryer). Linguistic typology generates theories about the relationship between different parameters of the grammar of a language and of languages, or about the possible range of human languages. Although formal linguistics in the Chomskyan tradition puts forward theories about the nature and organisation of the human language capacity or 'organ', the accompanying formal metalanguage is of course only a theory in a rather limited sense: either in the same sense that a descriptive grammar is a theory, or in the sense that its generative possibili-

ties and limits are supposed to mirror, directly, the exact range of human linguistic diversity.

In fact, the broadest theories on the nature of human language and languages require all three enterprises for their proper formulation and evaluation, along with many other subdisciplines such as historical linguistics, psycholinguistics, pragmatics, and so forth.

Starting from Chomsky's (1957) historically important but conceptually fraught distinction between *descriptive* and *explanatory adequacy*, the generative tradition has tended to promulgate a particular linkage of one class of formal models with the need to connect linguistics to deep general questions about the nature of language and what it implies about human cognitive capacity. This has conflated the jobs of formal modelling and explanation for whatever universals the study of human languages may reveal. The 'explanation' then becomes simply the most elegant model – the model which best accounts for the sanitized, cocooned 'fragment' of a language or set of received grammatical test cases. To the extent that this occurs, generative linguistics neglects wider questions of empirical/descriptive adequacy. It does not, as a matter of course, rigorously test its models against the full complexity of individual languages – presumably a job for descriptive linguistics with regard to specific languages⁴ – or assess empirically its claims of language universals against the true diversity of the world's languages – presumably a job for typologists (see Newmeyer 2005).

Of course grammarians of individual languages aspire to descriptive adequacy for their grammars, but they also seek to explain the patterns they find. While this may include an appeal to features attributed to a purported 'Universal Grammar', formulations of this type can explain only a small part of what we recognise as the special characteristics of an individual language. Explanations for patterns in particular languages may lie outside the individual speaker's linguistic competence, in the history of the language, in its cultural matrix, in the wider sociolinguistic context in which the given language is only one of a number of codes in use, or in those aspects of the use of the language which might be described as lying in the realm of 'performance' rather than 'competence.' Moreover, the successful search for explanatory adequacy, because it targets the language capacity in general rather than the particularities of any one language, requires us to synthesize general principles of linguistic organization from the wide variety of linguistic possibilities that the enterprise of linguistic typology lays out for us.

We briefly illustrate this division of labour by discussing the different angles that each of these three enterprises would take on the problem of 'active/stative' languages where the arguments of different classes of intransitive verb behave quite differently in terms of their syntactic properties, in some cases grouping with the patients of bivalent verbs, in others with the agents of bivalent verbs, and possibly exemplifying other behaviours as well (e.g. grouping with the beneficiaries of three-place verbs).

The challenge to the descriptive grammarian is to track this behaviour through as many parts of the grammar as possible, e.g. pronominal agreement, behaviour of various subordinate clause types, binding of quantifiers, interaction with topicalization, and so on. Since such languages are often best described without reference to notions of 'subject' and 'object' (because of the splitting of subject properties) the descriptive grammarian must also develop an appropriate general framework of grammatical relations suitable for the languages (perhaps drawing on particular formal models, or on concepts available in typology). For example, Durie's (1985) analysis of Acehnese makes use of actor and undergoer macro-roles,⁵ which are carried through all relevant parts of the description. Additionally, the grammarian must also specify which verbs belong to which class, steering the right course between general principles (e.g. 'non-volitional intransitive verbs have patient-like arguments') and appropriate detailed lexical particulars (a listing of which verbs go in each class).

The challenge to the typologist is to synthesize the similarities and differences across 'role-based', 'active' or split-S languages such as Acehnese, Guaraní, Tsova-Tush and so forth, along the full range of relevant dimensions. How many roles are there? Which construction types are the splits manifest in? What semantic properties can one identify as relevant in affected verbs? How do the properties correlate? Can one find implicational hierarchies that might be revealing of the universal properties of such constructions in human language? As the typologist's syntheses of such questions snowballs over different languages, descriptions of individual languages may be pushed into new territory as individual grammarians realize they had neglected particular issues in their descriptions.

For the formal linguist, the challenge is to develop a representational system that elegantly models the properties described in the descriptive grammar. Should they be captured by a particular type of phrase structure representation (e.g. by representing patient-type intransitives as only having an object at some level, the so-called 'unaccusative' solution), by a formalism employing macro-roles, or by some other method. Because of

their power to model correlated phenomena, debates between different formal models often also throw back questions to the descriptivist to investigate and describe in more detail (e.g. the interaction of different verb classes with floated quantifiers or complement types).

We believe that a proper understanding of these respective domains of linguistic inquiry can only help the field. Our goal in this book, however, is more focussed: to examine issues of descriptive grammar-writing against this broader backdrop.

3. Dimensions of description

3.1. Grammar within the Boasian trilogy

The modern classic division of labour in descriptive linguistics – the so-called ‘Boasian trilogy’ – envisages a documentary trilogy containing grammar, dictionary and texts, between them amounting to a fairly complete portrait of the language, and integrated so as to be mutually consistent in analysis, terminology, and the variety and range of language they account for. Because the contribution of descriptive grammars to ‘catching language’ can best be evaluated as part of such a trilogy, it is worth saying a bit more here about each component. We will also return to this in the concluding part of this chapter.

3.1.1. Texts

The text collection seeks to show the language as it really is, and among other things provides a corpus against which the grammar’s claims can be tested, and which subsequent linguists may scrutinize for generalizations overlooked by the original grammarian. As Mosel points out in her chapter, the fundamental priority of textual material has not been widely acknowledged or appreciated within academia, resulting in asymmetries both in what gets published and what is acceptable as a doctoral thesis topic.

Though the scientific and documentary value of a rich and varied corpus is beyond dispute, grammarians differ greatly in how they see the relation of corpus to grammar.

At one extreme, a grammar may be almost completely oriented to accounting for all and only the material in the corpus: for a famous example see Heath’s Nunggubuyu trilogy – Heath (1981, 1982, 1984).⁶ Ideally, this approach entails a sophisticated method for linking statements in the grammar to textual occurrences – something handled in Heath’s grammar by long lists of cross-references sometimes occupying pages, but now implementable less clumsily through hypertext linkages. While this orientation to the corpus is not essential for the grammar, it remains the case that a body of textual material provides the evidence against which the claims made by the grammarian can be independently verified, at least within the bounds of the corpus.

The view, however, that grammars should be answerable just to a published corpus seems an extreme position in practical terms. It is common to hear constructions ‘on the fly’, in ordinary language use, that hardly make it into the traditional corpus, but are nonetheless an important and interesting part of the language. While modern documentary techniques make it easier to capture such data, we should still allow the serendipitous off-record observations of the linguist to count as valid sources of information, as summed up in Sutton’s (1978: xvii) observation that ‘my main “research tool”, if it can be called that, was paying attention’.

Limiting what the grammar should account for to a corpus also overlooks the fact that speakers may have quite clear and revealing judgments (including unacceptability judgements) on complex constructions that rarely if ever appear in a traditional textual corpus. This position has of course been stressed in the generative tradition since the late 1950s, and is persuasively restated here, for the purposes of descriptive grammar-writing, in Rice’s chapter on Slave. More pragmatically, time constraints on the careful transcription of data typically mean that linguists can proceed much more productively and interestingly with their overall grammatical analysis if they are not tied at every stage to a transcription and analysis of their complete recorded corpus. This makes for a more ‘spiral’ view of how documentation and description inform one another (Grinevald 2001: 288) – that is, advances on each of the three broad Boasian fronts interact with and build on advances on the other.

It may also be the case – see Hill’s chapter on the American tradition, and Donaldson’s (1980) grammar of Ngiyambaa for a nice example – that grammars may reflect a painstakingly curated choice of example sentences for their material, in such a way as to portray the culture of the speech community. A more principled way to do this, that linguists again have

barely begun to contemplate,⁷ is through the idea of an 'ethnencyclopaedia' – a representation of the full spectrum of a speech community's knowledge of the world, as expressed in the mother tongue. Though an immense undertaking, this certainly increases the chance that a textual corpus will successfully sample all relevant constructions in a language.

Recent developments in technology enable audio and video recording of language in use in ways that were not possible in the recent past. Thus the text component of the Boasian trilogy is typically understood as a collection of narrative text – a genre that may be relatively fixed in its performance and is at least more amenable to dictation than the sorts of ordinary conversation we can now record and transcribe. Boas himself was already well aware of the limitations this imposed, though he saw it as a limit on the syntactic elaboration of narrative rather than the availability of dyadic conversation:

The slowness of dictation that is necessary for recording texts makes it difficult for the narrator to employ that freedom of diction that belongs to the well-told tale, and consequently an unnatural simplicity of syntax prevails in most of the dictated texts. (Boas 1917: 1)

Developments in technology have changed the nature of the corpora we work with. As we accommodate these changes in our descriptive practice, our expectations of the grammars we write and about how the supporting evidence is to be provided will also change. We return to this issue in the final section of this chapter.

3.1.2. Dictionary

The dictionary, the second member of the Boasian trilogy, is normally the part of a linguist's documentation task that takes the longest to reach fruition. Gestation times of forty years or more are not uncommon. This largely reflects the diminishing-returns effect that follows from the statistics of language sampling. Data relevant to most generalizations about grammar surface regularly, so that after a couple of years' work on a language a field linguist generally has the feeling that they have encountered most grammatical patterns needing description – though the greater the scope expected of a grammar, the longer this point takes to reach, and expected coverage is growing all the time (Cristofaro, Rice). On the other hand, some lexical items may only surface once in many years, and only in very

specific conditions, particularly if the lack of equivalents in other languages means the investigator would not naturally think to ask about them. A typical example is the Kayardild word *marirrnga* 'single cooked *jardiyali* fish ["month-fish"] buried in sand by camp-fire to ensure return of a new school of fish next evening', which one of us (Evans) only discovered fortuitously after many years of working on the language.

Overall, there has been a relative shifting of effort away from dictionaries and towards grammars in the last half-century. Grinevald (2001) cites a contemporary 10:3:1 ratio in the publishing of grammars, dictionaries and texts. But this was certainly not the case in the nineteenth and early twentieth centuries, as witnessed by the many little-known languages (Mundari and Makassarese being two examples)⁸ that have exemplary dictionaries from the first half of the twentieth century but as yet no grammatical description of comparable scope.

In describing little-known languages, moreover, dictionaries suffer from the same 'decoding' bias that affects grammars (on which more below). Overwhelmingly they are one-way bilingual, from the target language to a national or world language (perhaps with a reverse finder-list), rather than monolingual, or symmetrically bilingual. The biasing effects this has on our documentation of these languages have yet to be fully appreciated, but we should certainly expand our notion of ideal documentation to include monolingual and genuinely bilingual dictionaries, as well as thesauruses.

The boundary between lexicon and grammar is a difficult one to draw in a principled way, and will inevitably involve a good deal of overlap, a problem that has long been clear to descriptive grammarians and language teachers, and which is now being more seriously addressed within general linguistics (Pawley and Syder, 1983, Fillmore, Kay and O'Connor, 1988, Kay and Fillmore, 1999, Croft, 2001). For example, in its treatment of clause-level syntax a grammar needs to contain lists of verbs with different valence patterns (see e.g. the long lists in Chapter 15 of Haspelmath's (1993) grammar of Lezgian). At the same time, each corresponding verb in the dictionary must indicate the available valence patterns. A dictionary should also include extensive information about the part of speech of each word, its patterns of government, gender, or the type of complement(s) it may take. The definitions, morphosyntactic behaviour, tests and argumentation for each of these categories, will normally be given in the grammar rather than the dictionary, but the two treatments need to be rigorously compatible and in that sense form part of an integrated larger work.

One example of this interaction comes from the problem of semi-regular 'heterosemy' – the use of signs with the same signifier and related meanings in different syntactic combinations – in languages that recycle verbs as prepositions. If a grammarian writes that 'verbs can also function as prepositions', using examples like *X take knife cut meat* for 'X cuts the meat with a knife' or *X fly come Beijing* for 'X flies to Beijing', is this a problem for the grammar or the lexicon? Balancing generalization with the right level of specificity is a tricky analytic problem, the subject of Enfield's chapter.

Another particularly fraught problem at the boundary of lexicon and grammar involves complex verbal expressions in which a single lexeme manifests as two distinct grammatical words, a light verb (or auxiliary, or function verb) and a second element (illustrated by the phrase type *give a shove*). Though rather peripheral in English, such constructions are crucially important in many languages of Africa, Eurasia, New Guinea and Australia. Two chapters of this book (Schultze-Berndt, Amha and Dimmendaal) are devoted to the problems of description and cross-linguistic comparability that these constructions pose. They show clearly how distinct descriptive traditions are as important as the facts of the individual languages themselves in preventing the development of a genuinely integrated cross-linguistic terminology.

3.1.3. Grammar

Turning our focus back to the grammar, there are many factors that shape the overall conception and organization of a grammatical description: linguistic variety, relation of metalanguage to described language, prescriptive versus descriptive function, audience (linguist, teacher, member of speech community), goal (reference grammar, pedagogical grammar) and direction (analytic, from form to meaning; or synthetic, from meaning to form).

Mosel's chapter investigates these dimensions in detail. In general, the grammars described in this volume are (to use her categories) grammars of spoken vernaculars, where the metalanguage is a language of wider communication (typically a colonial or other world language), and written for linguists to use as reference grammars.⁹ We focus on these because we believe that, even where a community or linguist hopes to produce other sorts of grammar (e.g. a monolingual grammar for use in the speech community, or a pedagogical grammar), the particular configuration we have outlined forms the most natural first step in the process of analysing a new

language and helps set up a framework which can be drawn on for other purposes. Two fine examples of reference grammars which succeed in reaching out to a broader audience (particularly the community of native speakers) through judicious structuring and careful explanation of terminology are Bauer's (1997) grammar of Maori and Valentine's (2002) grammar of Nishnaabemwin.

Within the broad characterization given above, one asymmetry is particularly important: that between analytic / decoding / semasiological / form-based grammars, that take as their starting point forms or constructions in the target language, and synthetic / encoding / onomasiological / meaning-based grammars, that start from particular meaning categories (e.g. tense, or space, or causality) and show how grammars – in conjunction with the lexicon, where necessary – allow meanings within these fields to be expressed.

The ideal of having both types of grammar for any given language goes back to Gabelentz (1891), and Mosel's chapter discusses the history of this bidirectional conception, and the problems faced in seeking to implement it. But for many reasons comprehensive meaning-based grammars have hardly ever been produced, with the honourable exception of Leech and Svartvik's (1975) 'communicative grammar of English', although some descriptive grammars – such as Seiler (1977) on Cahuilla or Wilkins (1989) on Mparntwe Arrernte – contain a sample chapter or chapters written from a meaning-based perspective, and there have been other interesting experiments in grammatical organization, such as Newman's 'encyclopaedic reference grammar' of Hausa, that includes some entries based on form and others based on meaning. Despite the difficulties in executing this ideal, however, we believe this should remain an important descriptive goal. Arguably only in taking a meaning-to-form based approach to grammatical analysis can we hope to model the process by which a speaker of the language encodes their thoughts in speech. Additionally, many of the questions asked by linguistic typology approach language comparison from the standpoint of meaning or function rather than form (Cristofaro, Zaefferer).

One of the crucial problems faced in writing synthetic grammars is that the techniques of structural linguistics have always been more successful at dealing with form than with meaning. Form is also something that is easier for non-native speakers to grasp from an early stage of research than meaning is, which makes the analytic approach sit more naturally with the limited knowledge of non-native speaker linguists, who parallel other second language learners in being far more advanced in their passive than their

active grasp of the language. The production of good meaning-based grammars is thus an area where the involvement of native-speaker linguists is particularly important. Ameka's chapter discusses the difference between his facility in generating examples that require encoding ability, when working on Ewe as a native-speaker linguist and on Likpe as a non-native speaker linguist.¹⁰

Recent advances in semantic fieldwork, however, have begun to give us better techniques for tackling these problems: see Hellwig's chapter for an application of 'field semantic' techniques to the description of location, posture and aspect in Goemai. Further, developments in hypertext (Zaefferer) lay the technical foundation for the eventual construction of electronic grammars where the reader can access material from either an analytic or a synthetic perspective, though this still remains a dream rather than a reality.

3.2. Connecting with context; culture, history and heterogeneity

In addition to delineation problems between grammar and dictionary, grammar-writers face other tricky questions about where to draw the boundaries of their description. However much 'an inclusive scope may ... be at odds with professional pressures to build up an elegant, concise and typologically unified grammatical construct' (Diller), this entails a sacrifice of realism that most grammarians are unwilling to make, given that '[f]ield research ... cannot find either a homogenous speech community or an ideal speaker-hearer' (Heeschen 1998: 39). Within the history of linguistics, one current that runs from Panini to Chomsky has striven for formal elegance by treating language as an autonomous entity, by holding 'the view that all connections between language and non-linguistic matters can be disregarded'¹¹ (Anttila 1973: 177). Pulling against this is another current, associated traditionally with the philological quest to understand and interpret language in its cultural context, that aims 'to capture precisely those ways in which language is linked to knowledge of the world' (Hill 1989: 119), and which heeds Hockett's dictum that

Linguistics without anthropology is sterile, anthropology without linguistics is blind. (Hockett 1973: 675)

Linguists sympathetic to this more ethnographic and philological approach need to confront three particularly important boundary questions in

determining the scope of their descriptions. What are the boundaries between language and other aspects of culture (3.2.1)? How does one handle the relationship between the contemporary synchronic state and previous *états de langue* (3.2.2)? And how does one steer a course between describing the variety representing an idealized average speaker (of a particular lectal variety) and the complex range of varieties considered by the speech community, or the linguist, to function as an integrated system?¹² The diglossia resulting from certain types of language contact may result in situations like that described in Diller's chapter on Thai, where:

higher and lower registers look somewhat different typologically. Common Thai is mainly monosyllabic and isolating, forming compounds on a head-first principle. ... [R]oyal Thai is highly polysyllabic and derivationally agglutinating, often forming head-final compounds.

The three boundary questions mentioned above, in their turn, grade into one another. A prestige variety of a language may reflect an earlier historical reality (Joseph), or the influence of another prestige language, perhaps in a locally-modified version (Diller). Cross-dialectal variation may provide information about the path of language change. Or cultural factors may shape the interpretation of lectal variation,¹³ such as the famous case of 'moiety lects' in Arnhem Land where the variation is geographical and temporal (one lectal variant being more conservative), but is emically interpreted in cultural terms (with variants linked to culturally-construed patri-moieties).

In each of these cases, there is a tension between autonomist, essentialist approaches, which focus on an idealized homogenous variant with the goal of developing the most elegant and concise account, and more inclusive approaches which admit variation to the description. The reasons for taking the latter course may include a wish to depict realistically the semiotics of the full speech community, or to document variation which may in turn give a better understanding of the underlying logic of the linguistic systems under analysis, or provide materials for analysing change in progress, including evidence of on-going processes of language contact. In general, autonomist approaches treat grammar as (idealized) individual knowledge, and language as a distinct modular system, while non-autonomist approaches treat language as a social phenomenon, and problematize the separation of language from other elements of culture.

3.2.1. *Grammar and culture*

Despite the genesis of the Boas-Sapir descriptive tradition in anthropology, and a widespread perception that this represented an ideal of culturally-embedded language description, grammar-writing in this tradition in fact paid far less attention to the interactions between grammatical categories and culture-specific cognitive categories than is normally believed (Hill). To some extent we might see this as an outcome of the Boasian trilogy and the division of labour it plots. The embedding of language in culture is made evident in the selection of culturally important texts to illustrate the narrative genius of the language. The dictionary too makes clear that language and culture are two sides of the one coin. Yet there are also many areas of grammar, such as the encoding of kinship, evidentiality, or intentionality, where culture-specific categories intrude (Enfield 2002), and there are good reasons to admit such categories into grammar-writing.

Firstly, as Hill points out in her chapter, if we are 'to base our case for documenting and developing threatened languages largely on a claim that they and their speakers contribute irreproducible understandings to the total store of human knowledge, we must do far more work of this type'.

Secondly, revival of interest in the Humboldt-Sapir-Whorf question of the nature of the relationship between language and thought (see e.g. Levinson 2001, Lucy 1997) promotes culture-specific categories to important test-cases that can be used as independent variables in examining the influence of language on cognition.

Thirdly, the recognition of the importance of complex culture-specific categories in grammar is directly relevant to debates about the extent to which fundamental grammatical categories are innate and to what extent these can be mastered by learning children without preexisting concepts to hook grammatical categories to. The more the categories found in a given grammar are culture-specific – such as the kinship-sensitive pronoun terms and derivational verb forms based on generational moieties common in Australian languages (see Dench 1982, 1987 for complex examples) – the less likely it is that the child brings innate, preexisting knowledge of them to the task of learning their language.

Fourthly, the existence of such culture-specific categories provides evidence directly relevant to debates on the relative influence of universal vs culture-specific factors on the development of linguistic structures, as well as the question of how far language change is completely blind and how far

it responds to linguistic ideologies, albeit in unpredictable ways (Silverstein 1985, Simpson 2002, Evans 2003a).

3.2.2. *Grammar and history*

The Saussurean prohibition against mixing synchrony and diachrony, bolstered by the Chomskyan argument that the language-learning child must construct their grammar without reference to anything but synchronic facts, have contributed to a view that grammars should give purely synchronic accounts of the languages they describe.

No descriptive grammarian today would want to include diachrony to the extent found in some nineteenth century grammars that appear almost entirely interested in the historical origins of grammatical formatives – or, conversely, only interested in formatives that descend from elements known in older structures. However, there are good reasons to include diachronic discussions within a grammatical description, as discussed at greater length in the chapters by Rankin, Joseph and Chappell.

First, it has been said that every language is something of a ruin, and often the best explanation that can be provided for a particular grammatical pattern is a historical one. Rankin makes this case very clearly in his chapter. A satisfying description generally succeeds in providing a natural characterization of the conditions or environments across which different allomorphs are distributed. Accounts that fail to do this, and appeal to 'diacritic' or unmotivated rote-based conditioning, are normally viewed as unsatisfactory, or as having failed to solve the descriptive problem. Rankin discusses a Siouan case study of verb-stem ablaut for which no satisfactory synchronic account had been found by generations of scholars, and where diachronic evidence shows why the seemingly unsatisfactory rote-based account is in fact the correct one, reflecting the outcome of a particular sequence of analogical changes that has destroyed the original phonological basis for the alternation. Explanations for patterns in the grammar, then, cannot always be found in language universals, and instead often lie in the idiosyncratic history of the particular language.

Second, as soon as one tries to represent lectal diversity in a speech community, the possibility arises that lectal differences are due to the coexistence of innovative forms or structures with older varieties maintained in literate or high registers. Joseph's chapter discusses Greek as an extreme case of this, where the *katharevousa* variety results from a complex process

of consciously archaizing movements at various points in the language's history importing norms from earlier classical stages. It can, of course, be difficult to draw the exact boundaries of what counts as an archaic register of the contemporary language, as opposed to an earlier variety that no longer exerts influence. But it should not automatically be assumed that older varieties are mere fossils, and Joseph gives as an example the modern importation of the inflected pronoun strategy for forming relative clauses from the learned *katharevousa* variety into standard modern Greek usage.

Apart from the particular value of taking prestige-variety sediments from older varieties into account in explaining synchronic variation in grammatical rules, there are broader implications for our understanding of language learning. Even if the young child only has access to what they hear of the synchronic state, it is naïve to assume that older children and even adults, as part of the enculturation process, are not influenced by overt education which may give a second track of transmission. How far and in what circumstances the more conscious schooling of older learners impacts on language structure remains an open question, and grammars that bring such cases to our attention provide valuable case studies. It is also wrong to represent more conscious language learning only as a phenomenon of large literate societies, since many non-literate societies also have means (such as the later learning of oratorical registers) for inculcating material from earlier varieties of the language.

There is a final reason to include diachronic material in grammars. Though tastes vary in how far grammars should discuss alternative analyses of the data presented, as opposed to merely stipulating a single satisfactory analysis, there are many places where it makes sense to consider competing analyses of particular phenomena, particularly when the grammarian is proposing an analysis that seems unusual or deviant compared to the first-blush analysis that an educated linguistic reader would assume for the data, or to analyses that are widely accepted for seemingly comparable constructions in other languages. It is often the case that a *prima-facie* analysis being rejected by the grammarian does in fact well describe an earlier phase of the language, but that subsequent changes in the language's organization leave this analysis as unsatisfactory. Treating diachronic change and argumentation about the best analysis together can simultaneously help convey the special nature of the phenomenon under discussion, and make the discussion of analytic alternatives more lively and less casuistical than it would otherwise be.¹⁴

3.2.3. Dealing with lectal variation

Virtually all speech communities contain at least some lectal variation so grammarians need to confront the fact of variation, except in the special case of languages based on a single speaker, where language death shrinks lectal variation to something like Chomsky's (1965: 3) 'ideal speaker-listener' in a 'completely homogeneous speech-community'. How to address the issue of lectal variation in a principled way, and how far to cast the descriptive net, is a difficult problem to which grammarians adopt widely-varying solutions, ranging from grammars focussing on just a single variety (e.g. 'grammar of the X dialect of Y'), to pan-dialectal or diasystemic grammars. Variation, of course, may reflect many factors – levels of formality (including oratory, learned registers and so forth), geographical variation, factors of gender, ethnicity or class, other constructed cultural categories like moieties (Morphy 1977) or clans (Sutton 1978), interpersonal relations (e.g. kin relations), or features of the referent or speaker (such as for the Royal Thai discussed in Diller's chapter).

Important foundations towards the development of polylectal grammars were laid by Fries and Pike's (1949) study of coexistent phonemic systems and Klima's (1964) transition rules linking different varieties derived from a single underlying structure. Weinreich's (1954) seminal article on diasystems showed how a single system can be discovered under dialect variants and so demonstrated that structuralist principles need not be restricted to a single variety. Labov's work shows that it is the speech community rather than the individual speaker where one can best find systematicity in variation, thus moving the goal of a grammarian who takes variation seriously towards Saussure's socially-based *langue* and away from Chomsky's individual-based competence.

Particularly within studies of creole continua, there have been attempts to construct polylectal grammars consisting of families of rules combining in highly constrained ways, and subject to implications as one moves up and down the spectrum from acrolect to basilect (DeCamp 1971, Bailey 1973). Despite these theoretical advances, the field has been slow to move beyond the sketching of polylectal grammar fragments to the production of complete grammatical descriptions. We have not yet moved as far as we might wish¹⁵ from the situation described by Derek Bickerton in 1973:

[T]hose who would write polylectal grammars are immediately faced with a major difficulty. Since extant grammars are all monolectal, they can have only the vaguest notions as to what are the properties and structure of a

polylectal grammar, and which of a number of possible forms it might best take. (Bickerton 1973: 18)

Despite the clear challenges to be faced in actually writing such a grammar, there are many reasons to see it as a desirable goal. Taking the repertoire of a multi-lectal speech community as the relevant object for description can help to make sense of patterns which might otherwise make little sense in individual lects. For example, pan-lectal comparison in the Bininj Gun-wok dialect chain (Evans 2003b) allows a better understanding of the processes of neutralization that create apparent asymmetries in some dialects in the systems of gender agreement on the one hand, and verbal agreement for subject and object agreement on the other. Thus, polylectal grammars may play a crucial rule in helping us to understand transitions between different synchronic states. Studies of creole continua, or of diglossic systems like Thai (Diller), are crucial in showing us how the effects of language contact are played out through lectal variation.

Though detailed dialectological work properly belongs elsewhere than in a grammatical description, there are places where a grammatical description that succinctly ties together cross-dialectal evidence can take advantage of 'bountiful nature [which] has given us three great linguistic laboratories: diachronic, synchronic and dialectology' (Moulton 1969: 460). A nice illustration of this is Hyman and Tadjadjeu's (1976) demonstration of how different dialects of the Bamileke language illuminate the emergence of 'floating tone': in conservative dialects, all tones are linked to segments, while more innovative dialects have lost the segmental material but conserve the tones previously linked to them. As Hyman (2001: 26) puts it: 'it's harder for a group of dialects to hide their "stories" ... [G]et related languages. They will tell related stories.'

Language change is the point at which function shapes and reshapes structure, and thus the locus for recasting the question 'why are languages the way they are' as 'what forces cause languages to change into what they are'. Because variation – whether across dialects, individuals, or styles – is the most sensitive measure of change in progress, and in fact it is widely assumed that most language changes pass through a phase where two states are copresent in the form of variation, documenting variation is an important way that grammatical descriptions ultimately contribute to explanations about the nature of language.

Another advantage of studying multilectal systems is that one lect or register may reveal insights about semantic generalizations that remain covert in another lect. A classic employment of this method is Dixon's

(1971) study of the 'mother-in-law' variety of Dyirbal, and similar techniques applied to Royal Thai are discussed in Diller's chapter. Though the many-to-one relations are most typically lexical, these take on special significance when investigating verb classes, which as we saw above is a particularly important area of intersection between grammar and lexicon. A classic example is the evidence presented in Hale (1971), on the basis of work with lexical correspondences between ordinary and initiation registers, for the emic status of perception verbs as a superordinate category in Warlpiri.

Perhaps most importantly of all, expanding one's descriptive brief to the broader speech community can make for more realistic modelling of what individual linguistic choices mean, since the grammatical choices made by speakers are designed not just to convey propositional meaning, but to convey social meaning as well, even though individual competence is 'a smaller subset of the polylectal grid for productive competence, and a larger one for receptive competence' (Mühlhäusler 1992). Similar points can be made for knowledge of varieties within a dialect continuum. Taking a polylectal perspective thus offers the promise of modelling, ultimately, individual repertoires; how far individuals can vary in their knowledge of different varieties, what the relations are between these varieties, and what are the sociolinguistic categories that get reflected in these varieties.

4. The future of grammar-writing

One part of the challenge of grammar writing has always lain in presenting a coherent and reasonably comprehensive description of a complex linguistic system within the limitations of a published volume. The other part of the challenge, as we have seen, lies in adjusting to changes in the field of linguistics and our evolving understanding of the world's languages and of the human language capability, and in the nature of the available corpora and the media in which our results might be presented. In this section we discuss three issues that are likely to figure increasingly in the practice of future grammarians.

4.1. Rethinking the Boasian trilogy

Recent advances in recording technology and digital analysis have implications for our conception of 'corpus of data'. We suggested in 3.1.1 that the traditional text collection, amongst its functions, provides data against which analyses presented in the grammar can be verified. That the traditional corpus consists largely of texts is in many respects a reflection of what have been until recently clear limitations in recording technology. Field recordings have consisted of dictated vocabulary, sentences and texts in written form with a few selected audio recordings, again typically of narrative texts. Only in the last decade or two has it become standard to make audio recordings of all or most elicitation sessions held in remote locations. Today the relative ease of video as well as audio recording makes the more systematic collection of conversational data and of authentic language use a real possibility for field linguists. This makes possible the successful and detailed analysis of gesture, turn taking cues in gaze and kinesethics, external context etc., in ways that were not previously available except in specially arranged sessions. At the same time, software packages allow acoustic phonetic tagging and analysis of recorded materials that had until recently been possible only after painstaking work in specialised, and expensive, laboratories.

Digital audio and video recording, portable storage, and the development of software enabling the tagging, management and analysis of collected data raises the stakes for corpus collections. Our traditional published text collection consisted of a few hundred pages of narrative text with interlinear glosses, free translation and explanatory notes, but the modern published corpus may potentially consist of digital audio recordings of data collection sessions, some with accompanying video, and linked to a range of transcriptions representing different kinds and levels of analysis. Where the published text collection once served as the grounding evidence for a linguistic analysis, the digital archive will come increasingly to fill that role.

Recently, a number of writers (most clearly Himmelmann 1998, 2006, and Lehmann n.d.) have made a strong case for a distinction to be drawn between descriptive linguistics and documentary linguistics, with the latter having an agenda independent of that of the former. Traditionally, the nature of the corpus collected for living languages has been determined by the concerns of descriptive linguistics. But this is clearly changing. The need to define a separate endeavour of documentary linguistics comes not only

from the recognition of the severe limitations of a documentation program dependent on the concerns of the descriptive linguist, but also because the opportunities for broader documentation, given new technology, have increased markedly and there is at the same time a critical need for such broad documentation in the face of mass language extinction. What does this mean for the grammar writer?

We should anticipate that the expansion of the corpus to include real language use will effect a corresponding expansion in the grammar. Where most grammars have, to date, concentrated on constructions found in narrative – often complex sentence constructions more likely to surface in practiced and refined traditional narrative art – we can anticipate a shift in balance towards the linguistic forms more often found in naturally occurring everyday dialogue, and which have been more often emphasized in grammars written for pedagogical purposes. This shift is also consistent with some shifts in emphasis within the field of linguistics as a whole, as we discuss further below, and is likely to nudge a 'naturalistic turn' in the sort of data grammarians focus on accounting for.

The collection of narrative texts and the dictionary traditionally accompanying the grammar will instead or as well be presented as a collection of digital files including written versions, audio and sometimes video recordings. Of course, with current technology, the production of such material on disk is cheaper for a publisher, and for the purchaser, than an accompanying volume of texts and we can expect that the increasing use of electronic corpora will increasingly fuel demand for the grammatical description also to be presented as a hypertext document – more on this below.

Technological optimists should not forget, however, that the increased speed of producing digitised recordings is unlikely to be paralleled by a significant acceleration in how long it takes field linguists to produce the sorts of careful translations and cross-questioning of semantic issues that are the hallmark of a well curated text collection (see Evans and Sasse 2003). This will limit the proportion of collected materials that can be made available with full annotations. Rather than a hugely increased bulk of carefully annotated texts, then, what the new technologies are more likely to deliver us is a set of Russian dolls – a core of well-analysed material that may not greatly exceed in quantity that produced by more traditional methods, with an expanded set of roughly transcribed material and a huge amount of raw data without significant transcription or translation.

As documentary linguistics develops an agenda distinct from that of descriptive linguistics the relationship between the grammar and the published corpus, at least in the case of minority languages, will also change. A grammatical description grounded in a corpus collected for the purposes of writing that grammar is a very different thing from a grammar grounded ultimately in a corpus collected for a range of sometimes very different purposes. It will be necessary for the grammarian to define much more clearly what the grammar is a description of, what parts of the corpus it relates to, and of the available language materials what may have been selectively ignored, or which have not yet been analysed to the point where they can be used to test the grammar's claims.

4.2. Directions in the wider field

The impacts of a technology that allows us to build and investigate richer corpora are found not only in the study of minority languages. Where the focus of formal linguistics in much of the past 40 years has been on the competence of the idealized native speaker-hearer and through this the characterization of the innate human language capacity, in something of a 'performance turn' there has been a growing understanding of the need to build models of language use, processing and production, and many of these models are highly formalized. In part, this is a response to developments in natural language processing and the practical concerns of managing effective machine-human language interaction. To know how to build a machine that uses language in a human-like way, you first need to know how humans use language. Investigation in this domain relies much more on large and representative corpora than on categorical judgments about the intuitive grammaticality of what may in practice be vanishingly rare constructions.

The shift in formal focus can be seen perhaps most clearly in recent developments in formal semantics.¹⁶ Where syntax has traditionally sought to model an autonomous component regulating the structure of linguistic strings, semantics, being concerned with the construction and apprehension of meaning, seeks to describe how such meanings are mapped onto linguistic forms. This concern has led in recent times to an increased interest in the formal description of linguistic objects larger than the sentence. Thus we see the development of formal accounts of discourse and conversation that seek to show not only how such larger units of language cohere, but how

meaning is constructed dynamically (Kamp and Reyle, 1993, Asher and Lascarides, 2003). Not surprisingly, we also see the development of 'functional' formal models of syntax which seek to take account of sentence production and processing in real time (e.g. Kempson, Cann and Marten, 2005).

Discourse oriented approaches have an inherent emphasis on the parsing and interpretation carried out by native hearers, rather than on the sentence generating competence of native speakers. It is only natural that such approaches should favour the collection of naturalistic conversation, over the formal elicitation of grammaticality judgements. Perhaps as a result of this turn towards naturalistic data and a concern to model human language interaction, those working within formal frameworks are looking beyond the traditional fragments of European language data against which many formal accounts of discourse and conversation have been traditionally tested. The successful transmission of meaning is dependent not only on (presumed) universals of human cognitive processing – a reliance on pragmatic inference is a fact of all human communication – but also on the structure of the linguistic code. It has long been recognized that the structure of the code is a function of the uses to which it is put: 'grammars do best what speakers do most' (Du Bois, 1985). To the extent that these functions are universal we can anticipate the identification of formal linguistic universals, but to discover these we need to look at a wide diversity of linguistic systems.

Thus, we might anticipate an increasing engagement between discourse and performance oriented branches of formal linguistics, descriptive grammars, and linguistic typology. The descriptive frameworks preferred by typologists have tended towards the functional end of what has often been described as a formal/functional split in the field – so called functional approaches generally providing a richer and more satisfying account of the range of possible linguistic systems than those in the mainstream generative traditions (and see Newmeyer 2005). However, as those engaged in formal modeling increasingly seek to account for performance effects both in language use and in the structure of the code, the formalist/functionalist dichotomy begins to dissolve.

What does this mean for the grammar writer? The new study of generative syntax in the 1960s and 1970s gave grammarians new questions to ask (Rice), new tools with which to analyze languages and an imperative to provide more detailed descriptions of clausal and interclausal syntax. In the same way, we can expect the development of rigorous formal models of

connected discourse to provide us with a host of new questions, with new tools for analysis, a range of possible frameworks for description, and a clear expectation that grammars include detailed descriptions of discourse across a range of genres. For example, early Discourse Representation Theory arose in part out of attempts to describe the semantics of particular tenses. Thus Kamp and Rohrer's (1983) account of the differences between the French *passé simple* and *imparfait* places a great deal of importance on the semantic relationship between sentences containing these tenses and the sentences preceding them. If the semantics of tense necessarily involves reference to discourse context – and Discourse Representation Theory makes a strong case for this – then there are clear implications for grammar writing. The semantics of one of the most fundamental sets of morphological categories, namely verbal TAM inflection, cannot be adequately described without reference to a well described supra-sentential discourse context. Formal approaches to describing such discourse contexts can then provide us with questions we might ask of any language, tools with which to ask, and a possible vocabulary with which to present the answers.

In the relatively recent past our syntactic descriptions have been mainly concerned to reveal a native speaker's competence. They have focused on what strings it is possible to generate, and what it is not possible to generate, often illustrating the argument with examples constructed especially for eliciting grammaticality judgments. In describing particular forms and constructions, we must now pay much greater attention to the frequency with which these occur in particular genres, the range of interpretations possible, and in general terms, as already mentioned, we need to provide a clear characterization of the corpus, or fragment of a corpus, from which the analysis is derived. Grammars will expand to include descriptions of different discourse genres as a matter of course, including accounts of conversational structure, of variability in the corpus and variation in usage.

4.3. New forms of grammars

As we have noted a number of times in the preceding discussions, and as is treated in detail in Zaefferer's chapter, our new technology makes possible alternative forms for published grammars – grammatical descriptions that embrace the new media, integrate audio-visual material and that are less encumbered by the inherent linearity of simple written text. We have suggested here that such grammars are probably inevitable given the develop-

ment of enriched digital corpora, but along with the benefits they promise in flexibility there are challenges to be faced.

Hypertext grammars can provide a wealth of possible arrangements of information about the grammar, ideally allowing users to select different degrees of detail and to navigate the description via any number of different paths. Given that digital files are inexpensive to update compared with grammars published in book form, they also provide a degree of temporal flexibility.

There are two obvious and clearly related practical challenges arising from this high level of flexibility – version control and internal analytical consistency.

The problem of version control is typically not one faced by the grammarian, though it is understood by any lexicographer. The dictionary of a language is never finished: there are always new words to add, new meanings to describe for words already listed, and reanalyses of previously listed forms to be presented. We expect dictionaries to be incomplete, to lag behind actual language usage, and to eventually appear in new editions. Our expectations of grammars, have been quite different – we expect relatively definitive statements of what it means to know and speak the language in question. For those grammarians engaged in describing a small-community language they will outlive, this has taken an extreme form, since it may be that no other linguist will have the chance to engage directly with the language intensively enough to reevaluate their description of how the language works.

The inherent flexibility of the hypertext grammar may encourage us to think of grammars as works in progress in much the same way as we think of dictionaries. But there is a critical difference between the two. Where it is possible within limits simply to add to a dictionary without major revisions to what already exists, and without compromising the consistency of the analyses contained within, this is almost never possible for a grammar. Any reanalysis of a subsystem of the grammar must be checked against other parts of the grammar to ensure that the larger analysis remains coherent. When writing grammars of minority languages, we have typically understood that there will be some stage at which the grammar is 'finished' and the analysis is expected to be consistent at just that point in time. This consistency extends to the parses provided for the accompanying texts and for the category labels and other analyses contained within the accompanying dictionary. That consistency must derive from the deepest characteristics of the language, is well illustrated by languages like Tagalog or Chi-

nese (see chapters by Himmelmann, Chappell and LaPolla) whose radically different organizing principles have meant that far-reaching rethinkings of descriptive approach have been necessary in the course of reaching satisfying grammars of these languages. Typically this has meant that new approaches to describing such languages are undertaken as radically new undertakings by different scholars, rather than mere updatings of existing structures.

The organizational and temporal flexibility of the hypertext grammar/multimedia corpus makes managing consistency in grammars a much more challenging prospect than it is in dictionaries. On the one hand there is a danger of inconsistency, on the other there is a danger that a definitive grammatical description associated with a particular corpus of data may never be completed.

A fully developed and evolving hypertext grammatical description may represent not only the fruits of detailed linguistic investigation of a corpus but also a high-level analytical key to that (perhaps similarly evolving) corpus. As digital corpora and their annotations become more complex and their organization more sophisticated, the boundary between the description and the documentation may become increasingly blurred. Thus where the traditional corpus, as language documentation, may have been defined by the descriptive agenda, there is the possibility that hypertext grammars linked to multi-modal corpora may come to be shaped instead by the documentary enterprise. For grammarians, there will be challenges and choices to make here. While the grammatical analyses must remain true to the corpus, there will be a challenge in maintaining the autonomy of the grammar as a description of the language. With too little autonomy, the various linguistic analyses of diverse subparts of the corpus may not cohere into a consistent description.

The chapters of this volume make clear the complexity of the task of grammar writing and the diversity of skills required of the grammar writer. We hope that those reading this book will not only develop an enriched appreciation of the grammarian's craft but will be inspired to make their own contributions to this important and constantly challenging field of linguistic investigation.¹⁷

Notes

1. We regard it as an open question as to how successful grammars have been that try and build the whole description around a few organizing principles, or whether they should confine such remarks to an initial orienting preface – such as the interesting discussion in the introduction to Seiler's (1977) Cahuilla grammar of the principles of 'the preponderance of the descriptive over the labelling principle', 'syntactic compression', and 'weak centralisation' – or treat them apart from the grammar proper, as in Heath's (1986) article on non-configurationality in Nunggubuyu (which does a lot to unlock the difficulties of his rather forbidding 1984 grammatical description), or Launey's (1994) thoughtful and revealing essay on omnipredicativity as an organising principle in Nahuatl.
2. In his classic grammars of Dyirbal (Dixon 1972) and Yidiny (Dixon 1977), Dixon deliberately split the presentation between a more basic descriptive part and a separate chapter on syntax written within a particular formal framework (basically an adaptation of Aspects-style generative grammar). Of this first such attempt, he wrote frankly (Dixon 1972: 125): 'This chapter attempts to interpret, generalise from, and explain the basic facts [presented in previous chapters – AD and NE]; the discussion is thus at a higher level of abstraction, and is more speculative and arguable'.
3. In a variant of the Boasian trilogy, Kibrik split his masterful description of Archi into a basically descriptive 'taxonomic grammar' (Kibrik 1977a) and a 'dynamic grammar' (Kibrik 1977b) written within the theoretical framework of 'Meaning<-->Text' linguistics.
4. See Gross (1979) and Sadock (1996) for two particularly trenchant critiques of the answerability to descriptive precision within the generative tradition.
5. Durie's notions of actor and undergoer macro-roles, of course, derive from a particular theory, Foley and Van Valin's (1984) Role and Reference Grammar, that lies somewhere between a formalized theory and a typologically-motivated theory of generalizable linguistic categories. However, since his grammar is in other respects not couched within any particular formal framework, we use it to illustrate the approach of a descriptive grammarian rather than of a formal linguist.
6. Though corpus material was in fact supplemented by elicitation to ensure that complex paradigms were filled out: particularly in morphologically complex languages with large irregular paradigms a purely corpus-based approach is almost certain to end up with substantial gaps.
7. Though ironically perhaps the closest thing we have to an ethnoencyclopaedia was produced in the sixteenth century by Fray Bernadino de Sahagún, arguably the father of modern descriptive linguistics as applied to exotic languages. This takes the form of an enormous and comprehensive bilingual text collec-

tion portraying virtually every aspect of traditional Aztec life, published around 1557 (see Dibble and Anderson 1970, Sahagún 1905). This collaborative work drew on a large team of trained native-speaker linguists working with Sahagún to gather material at a large number of sites, as well as employing older Nahuatl artists as illustrators.

8. For Mundari, see the epochal sixteen-volume *Encyclopaedia Mundarica* (effectively a dictionary with encyclopaedic excursions) by Hoffman and Van Emelen (1930–1979), and for Makassarese, the magnificent dictionary by Cense (1979, though based on work carried out much earlier). It is only more recently that we have had anything like a reasonable grammatical description of either language (Osada 1992, Jukes forthcoming).
9. This corresponds closely to Lehmann's (1980: 29) idea of a *linguistische Grammatik* (linguistic grammar), which he defines as 'eine Grammatik, die dem durchschnittlich informierten linguistischen Fachgenossen Aufschluß über das Funktionieren einer Sprache auf allen Ebenen ihres Systems erteilen soll. Sie muß ihm den Vergleich dieser Sprache mit anderen ermöglichen'. (A grammar, which should provide the average informed linguistic colleague information about the functioning of a language at all levels of its system, and which should enable him to make comparisons of this language with others).
10. This point is also well made by Lehmann (1980: 31): 'Zum zweiten kann man die analytische Betrachtung der Sprachfakten einen Zugang "von außen", die synthetische einen Zugang "von innen" nennen. Die Bedeutung dieser Ausdrücke gewinnt praktische Anschaulichkeit, wenn man an linguistische Beschreibungsansätze denkt: den Zugang von außen nimmt der Linguist, der die Objektsprache noch nicht kennt und – gleichsam ganz Hörer – zuerst eine Analyse machen muß; den Zugang von innen nimmt der Linguist, der – gleichsam ganz Sprecher – die Objektsprache beherrscht und ihre Ausdrücke daher synthetisieren kann.' (Secondly we can term the analytic treatment of linguistic facts an approach 'from outside', the synthetic treatment an approach 'from inside'. The practical significance of these expressions becomes clearer, when one thinks of principles of linguistic description: the approach from outside is that taken by a linguist who does not yet know the object language and, like the hearer, must first carry out an analysis, while the approach from inside is that taken by a linguist who, like the speaker, has a mastery of the object language and can therefore synthesize its expressions'.)
11. Original: 'daß man alle Zusammenhänge von Sprache mit Nicht-Sprachlichem ausklammern könne' (Anttila 1973: 177).
12. In situations of language contact or mixed languages these may include partially integrated elements of one or more other languages. See Igla (1996) for a description of a Romani dialect where there is a particularly acute problem of dealing with far-reaching integration of other languages, here Turkish and Greek.

13. They may also affect local interpretations of individual difference in competence. See the discussion of Duranti's work on social variation in ergative case use in Samoan in Hill's chapter, and also the interesting observations in Aikhenvald's grammar of Tariana (Aikhenvald 2003: 311) on how individual differences in evidential proficiency are interpreted.
14. In other cases there may be no evidence of diachronic change in the variety under description, but two alternative analyses may each account equally well for the data. Here, too, treating multiple analyses may be relevant in another way, by showing bridging contexts that are particularly favourable to certain types of change through reanalysis.
15. For example, a 1992 encyclopaedia article by Peter Mühlhäusler on 'polylectal grammars' does not mention any comprehensive polylectal grammatical description.
16. For a discussion of the investigation of formal semantics in the field, see Matthewson (2004).
17. Just as this manuscript was going to press, a special issue of *Studies in Language* (Vol. 30: 2) came out, devoted to the topic *Perspectives on Grammar Writing*, edited by Thomas Payne and David Weber. It was too late to incorporate cross-references to this very interesting collection in the present book, but it contains a number of excellent papers addressing issues related to those dealt with here.

Acknowledgments

We gratefully acknowledge Felix Ameka, Nikolaus Himmelmann and Hans Kamp for their comments on an earlier version of this chapter, not all of which we have been wise enough to follow up on.

References

- Aikhenvald, Alexandra
2003 *A grammar of Tariana*. Cambridge: Cambridge University Press.
- Anttila, Raimo
1973 Linguistik und Philologie. In *Linguistik und Nachbarwissenschaften*, Renate Bartsch and Theo Vennemann (eds.), 177–191. Kronberg/Taunus: Scriptor.
- Asher, Nicholas and Alex Lascarides
2003 *Logics of Conversation*. Cambridge: Cambridge University Press.

34 Nicholas Evans and Alan Dench

- Bailey, Charles-James N.
1973 *Variation and Linguistic Theory*. Arlington, Virginia: Center for Applied Linguistics.
- Baker, Mark
1988 *Incorporation*. Chicago: Chicago University Press.
1995 *The Polysynthesis Parameter*. Oxford: Oxford University Press.
- Bauer, Winifred (with William Parker and Te Kareongawai Ewans)
1997 *The Reed reference grammar of Maori*. Auckland: Reed Publishing Ltd.
- Bickerton, Derek
1973 The structure of polylectal grammars. In *Sociolinguistics, Report of the 23rd Annual Round Table Meeting*, Roger W. Shuy (ed.), 17–42. Washington, D.C.: Georgetown University Press.
- Boas, Franz
1917 Introduction. *International Journal of American Linguistics* 1:1–8. Reprinted in *Franz Boas: Race, Language and Culture*, Franz Boas (ed.), 199–210. Chicago: The University of Chicago Press.
- Cense, A.A. (in collaboration with Abdoerrahim)
1979 *Makassaars-Nederlands Woordenboek*. 's Gravenhage: Martinus Nijhoff.
- Chomsky, Noam
1957 *Syntactic structures*. The Hague: Mouton.
1965 *Aspects of the theory of syntax*. Cambridge, Mass.: MIT Press.
- Croft, William
2001 *Radical construction grammar: syntactic theory in typological perspective*. Oxford: Oxford University Press.
- DeCamp, David
1971 Toward a generative analysis of a post-creole speech continuum. In *Pidginization and Creolization of Languages*, Dell Hymes (ed.), 349–370. London: Cambridge University Press.
- Dench, Alan
1982 Kinterms and pronouns of the Panyjima language of northwest Western Australia. *Anthropological Forum*, 5: 105–120.
1987 Kinship and collective activity in the Ngayarda languages of Australia. *Language in Society*, 16: 321–339.
- Dibble, Charles E. and Arthur J.O. Anderson (eds./translators)
1970 *Fray Bernadino de Sahagún, Florentine Codex*. University of Utah Press.
- Diller, Anthony
1993 Diglossic grammaticality in Thai. In *The role of theory in language description*, William A. Foley (ed.), 393–420. Berlin/New York: Mouton de Gruyter.
- Dixon, R.M.W.
1971 A method of linguistic description. In *Semantics: an interdisciplinary reader*, D. D. Steinberg and L.A. Jakobovits (eds.), 436–471. Cambridge: Cambridge University Press.
1972 *The Dyirbal language of North Queensland*. Cambridge: Cambridge University Press.
1977 *A grammar of Yidiñ*. Cambridge: Cambridge University Press.
- Donaldson, Tamsin
1980 *Ngiyambaa, the language of the Wangaaybuwan*. Cambridge: Cambridge University Press.
- Du Bois, John
1985 Competing motivations. In *Iconicity in syntax*, John Haiman (ed.), 343–365. Amsterdam: Benjamins.
- Durie, Mark
1985 *A grammar of Acehnese on the basis of a dialect of North Aceh*. Dordrecht: Foris.
- Enfield, N. J. (ed.)
2002 *Ethnosyntax*. Oxford: Oxford University Press.
- Evans, Nicholas
1995 *A grammar of Kayardild. With historical-comparative notes on Tangkic*. Berlin/New York: Mouton de Gruyter.
2003a Culture and structuration in the languages of Australia. *Annual Review of Anthropology* 32: 13–40.
2003b *Bininj Gun-wok: a pan-dialectal grammar of Mayali, Kunwinjku and Kune (2 vols)*. Canberra: Pacific Linguistics.
- Evans, Nicholas and Hans-Jürgen Sasse
2003 Searching for meaning in the library of Babel. Field semantics and problems of digital archiving. In *Proceedings of the Paradisec Workshop on Digital Archiving*, Linda Barwick and Jane Simpson (eds.). On-line publication available at http://www.paradisec.org.au/EvansSasse_paper_rev1.pdf
- Fillmore, Charles, Paul Kay and Mary O'Connor
1988 Regularity and idiomaticity in grammatical constructions: The case of *let alone*. *Language*, 64: 501–538.
- Foley, William A. and Robert Van Valin Jr.
1984 *Functional syntax and universal grammar*. Cambridge: Cambridge University Press.
- Fries, Charles C. and Kenneth L. Pike
1949 Co-existent phonemic systems. *Language* 25: 25–29.
- Gabelentz, Georg von der
1891 *Die Sprachwissenschaft. Ihre Aufgaben, Methoden und bisherigen Ergebnisse*. Leipzig.

- Grinevald, Colette
2001 Encounters at the brink: linguistic fieldwork among speakers of endangered languages. In *Lectures on Endangered Languages*: 2, Osamu Sakiyama (ed.), 285–314. Osaka: ELPR Publication Series 002.
- Gross, Maurice
1979 On the failure of generative grammar. *Language* 55: 859–885.
- Hale, Kenneth
1971 A note on a Walbiri tradition of antonymy. In *Semantics*, D. Steinberg and L. Jakobovits (eds.), 472–482. Cambridge: Cambridge University Press.
2001 Ulwa (Southern Sumu): the beginnings of a language research project. In *Linguistic Fieldwork*, Paul Newman and Martha Ratliff (eds.), 76–101. Cambridge: Cambridge University Press.
- Haspelmath, Martin
1993 *A grammar of Lezgian*. Berlin/New York: Mouton de Gruyter.
- Heath, Jeffrey
1981 *Nunggubuyu myths and ethnographic texts*. Canberra: Australian Institute of Aboriginal Studies.
1982 *Nunggubuyu dictionary*. Canberra: Australian Institute of Aboriginal Studies.
1984 *Functional grammar of Nunggubuyu*. Canberra: Australian Institute of Aboriginal Studies.
1986 Syntactic and lexical aspects of configurationality. *Natural Language and Linguistic Theory* 4: 375–408.
- Heeschen, Volker
1998 *An ethnographic grammar of the Eipo language*. Berlin: Dietrich Reimer Verlag.
- Hill, Jane
1989 The role of theory in language description. *Current Anthropology* 30: 119–123.
- Himmelman, Nikolaus
1998 Documentary linguistics and descriptive linguistics. *Linguistics* 36: 161–195.
2006 Language documentation: what is it and what is it good for? In *Essentials of language documentation*, Jost Gippert, Nikolaus P. Himmelman and Ulrike O. Mosel (eds.), 1–30. Berlin/New York: Mouton de Gruyter.
- Hockett, Charles
1973 *Man's place in nature*. New York: McGraw-Hill.
- Hoffmann, John and Arthur van Emelen
1930–1979 *Encyclopaedia Mundarica (XVI Vols)*. Patna, Bihar.
- Hyman, Larry M.
2001 Fieldwork as a state of mind. In *Linguistic Fieldwork*, Paul Newman and Martha Ratliff (eds.) 15–33. Cambridge: Cambridge University Press.
- Hyman, Larry M. and Maurice Tadadjeu
1976 Floating tones in Mbam-Nkam. In *Studies in Bantu Tonology*, Larry M. Hyman (ed.), 57–111. Los Angeles: University of Southern California.
- Igla, Birgit
1998 *Das Romani von Ajia Varvara. Deskriptive und historisch-vergleichende Darstellung eines Zigeunerdialekts*. Wiesbaden: Harrassowitz.
- Jukes, Anthony
f/c A grammar of Makassarese. Ph.D. diss., University of Melbourne.
- Kamp, Hans and Uwe Reyle
1993 *From discourse to logic: introduction to modeltheoretic semantics of natural language, formal logic and discourse representation theory*. Dordrecht: Kluwer.
- Kamp, Hans and Christian Rohrer
1983 Tense in texts. In *Meaning, use and interpretation of language*, Rainer Bäuerle, Christoph Schwarz and Armin von Stechow (eds.) 250–269. Berlin: de Gruyter.
- Kay, Paul and Fillmore, Charles
1999 Grammatical constructions and linguistic generalizations: The What's X doing Y? construction. *Language* 75: 1–34.
- Kempson, Ruth, Ronnie Cann and Lutz Marten
2005 *The Dynamics of Language*. Oxford: Elsevier.
- Kibrik, Alexandr Evgen'evič
1977a *Opyt strukturnogo opisanija arčinskogo jazyka. Tom II. Taksonomičeskaja grammatika*. Moscow: Isdatel'stvo Moskovskogo Universiteta.
1977b *Opyt strukturnogo opisanija arčinskogo jazyka. Tom III. Dinamičeskaja grammatika*. Moscow: Isdatel'stvo Moskovskogo Universiteta.
- Klima, Edward
1964 Relatedness between grammatical systems. *Language* 40: 1–20.
- Launey, Michel
1994 *Une grammaire omniprédicative. Essai sur la morphosyntaxe du nahuatl classique*. Paris: CNRS Édition.
- Leech, Geoffrey and Svartvik, Jan
1975 *A communicative grammar of English*. London: Longman.

- Lehmann, Christian
 1980 Aufbau einer Grammatik zwischen Sprachtypologie und Universalistik. In *Wege zur Universalienforschung*, Gunter Brettschneider and Christian Lehmann (eds), 29–37. Tübingen: Narr.
- 1989 Language description and general comparative grammar. In *Reference grammars and modern linguistic theory* Graustein, Gottfried and Leitner, Gerhard (eds.), 133–162. Tübingen: M. Niemeyer (Linguistische Arbeiten, 226).
- n.d. Language documentation: a program. Unpublished paper downloadable from www.uniefurt.de/sprachwissenschaft/personal/lehmann/CL_Publ/Language_documentation.pdf
- Levinson, Stephen
 2001 Covariation between spatial language and cognition, and its implications for language learning. In *Language acquisition and conceptual development*, Melissa Bowerman and Stephen Levinson (eds.), 566–588. Cambridge: Cambridge University Press.
- Lucy, John
 1997 Linguistic relativity. *Annual Review of Anthropology* 26: 291–312.
- Matthewson, Lisa
 2004 On the methodology of semantic fieldwork. *International Journal of American Linguistics* 70: 369–415.
- Mithun, Marianne
 2001 Who shapes the record: the speaker and the linguist. In *Linguistic Fieldwork*, Paul Newman and Martha Ratliff (eds.), 34–54. Cambridge: Cambridge University Press.
- Morphy, Frances
 1977 Language and moiety: sociolectal variation in a Yu:ngu language of North-East Arnhem Land. *Canberra Anthropology* 1.1: 51–60.
- Moulton, William G.
 1969 Structural dialectology. *Language* 44: 451–466.
- Mühlhäusler, Peter
 1992 Polylectal grammar. In *International Encyclopaedia of Linguistics*, Vol. 3, William Bright (ed.), 243–245. Oxford: Oxford University Press.
- Newman, Paul
 2000 *The Hausa language. An encyclopaedic reference grammar*. Yale: Yale University Press.
- Newmeyer, Frederick
 2005 *Possible and probable languages: a generative perspective on linguistic typology*. Oxford: Oxford University Press.
- Osada, Toshiki
 1992 *A reference grammar of Mundari*. Tokyo: Institute for the Languages and Cultures of Asia and Africa.
- Pawley, Andrew and Frances Syder
 1983 Two puzzles for linguistic theory: Nativelike selection and native-like fluency. In *Language and Communication*, Jack C. Richards and Richard W. Smith (eds), 191–225. London and New York: Longmans.
- Sadock, Jerry M.
 1996 PIFL: the Principle of Information-Free Linguistics. *CLS 32 (Parasession on Theory and Data in Linguistics)*: 133–138.
- Sahagún, Fray Bernardino de
 1905 *Historia general de las cosas de Nueva España*. Ed. Francisco del Paso y Troncoso (4 vols). Madrid: Fototipia de Hasuer y Menet. (Orig. 1557)
- Sapir, Edward
 1921/1949 *Language. An introduction to the study of speech*. New York: Harcourt, Brace and Co.
- Seiler, Hansjakob
 1977 *Cahuilla Grammar*. Banning, California: Malki Museum Press.
- Silverstein, Michael
 1985 Language and the culture of gender: at the intersection of culture, usage and ideology. In *Semiotic mediation: sociocultural and psychological perspectives*, Elizabeth Mertz and Richard J. Parmentier (eds.), 219–259. New York: Academic Press.
- Simpson, Jane
 2002 From common ground to syntactic construction: associated path in Warlpiri. In Enfield (ed.), 287–307.
- Sutton, Peter
 1978 Wik: Aboriginal society, territory and language at Cape Keerweer, Cape York Peninsula, Australia. Ph.D. diss., University of Queensland.
- Valentine, J. Randolph
 2001 *Nishnaabemwin Reference Grammar*. Toronto: University of Toronto Press.
- Weinreich, Uriel
 1954 Is a structural dialectology possible? *Word* 10: 388–400.
- Wilkins, David
 1989 Mparntwe Arernte (Aranda): studies in the structure and semantics of grammar. Australian National University: Unpublished Ph.D. Thesis.

Woodbury, Anthony

- 1993 A defense of the proposition, "When a language dies, a culture dies". In *SALSA I (Texas Linguistic Forum 33)*, Queen, Robin and Rusty Barret (eds.), 102-130. Austin, TX: University of Texas Department of Linguistics.
- 1998 Documenting rhetorical, aesthetic, and expressive loss in language shift. In *Endangered Languages: Current Issues and Future Prospects*, Lenore A. Grenoble and Lindsay J. Whaley (eds.), 2334-2358. Cambridge University Press.

Woolard, Kathryn, Bambi Schieffelin, and Paul Kroskrity (eds.)

- 1998 *Language Ideologies*. Oxford University Press.

List of contributors

Felix Ameka

African Linguistics
Leiden University
Postbus 9515
NL - 2300 RA Leiden
The Netherlands
F.K.Ameka@let.leidenuniv.nl

Azeb Amha

African Languages and Cultures
Van Wijkplaats 4
PO Box 9515m
2300 RA Leiden
The Netherlands
a.amha@let.leidenuniv.nl

Hilary Chappell

Centre de Recherches Linguistiques sur L'Asie Orientale,
École des Hautes Études en
Sciences Sociales
54 Bd Raspail 75270 cedex
Paris, France
Hilary.Chappell@ehess.fr

Sonia Cristofaro

Dipartimento de Linguistica
Università di Pavia
Corso Strada Nuova 65
I-27100 Pavia
Italy
russo@inrete.it

Alan Dench

Linguistics
University of Western Australia
Crawley WA 6009
Australia
Alan.Dench@uwa.edu.au

Tony Diller

National Thai Studies Centre,
Australian National University,
Canberra ACT 0200
Australia
Anthony.Diller@anu.edu.au

Gerrit Dimmendaal

Institut für Afrikanistik
Universität zu Köln
Meister-Ekkehart-Str. 7
D-50923 Köln
Germany
gerrit.dimmendaal@uni-koeln.de

Matthew Dryer

Department of Linguistics
University at Buffalo
Buffalo, NY 14260
USA
dryer@acsu.buffalo.edu

N. J. Enfield

Max Planck Institut für
Psycholinguistik
PB 310,
6500 AH Nijmegen
The Netherlands
Nick.Enfield@mpi.nl

Nicholas Evans

Linguistics
University of Melbourne
Parkville, Victoria 3010
Australia
nrde@unimelb.edu.au

Birgit Hellwig

Research Centre for Linguistic
Typology
LaTrobe University
Bundoora, Victoria 3086
Australia
bh21@soas.ac.uk

Jane Hill

Department of Anthropology
The University of Arizona
PO Box 210030
Tucson, Arizona 85721-0030
USA
jhill@U.Arizona.EDU

Nikolaus Himmelmann

Sprachwissenschaftliches Institut
Ruhr-Universität Bochum
D-44780 Bochum
Germany
himmelma@linguistics.rub.de

Brian Joseph

Department of Linguistics
The Ohio State University
Columbus, Ohio 43210-1298
USA
bjoseph@ling.ohio-state.edu

Randy LaPolla

Linguistics
LaTrobe University
Bundoora, Victoria 3086
Australia
R.LaPolla@latrobe.edu.au

Igor Mel'čuk

Department of Linguistics and
Translation
University of Montreal
Montreal, Quebec H3C 3J7
Canada
igor.melcuk@umontreal.ca

Ulrike Mosel

Allgemeine und Indogermanische
Sprachwissenschaft
Universität Kiel
D-24098 Kiel
Germany
umosel@linguistik.uni-kiel.de

Dory Poa

Research Centre for Linguistic
Typology
LaTrobe University
Bundoora, Victoria 3086
Australia
dorypoa@yahoo.com

Robert Rankin

Department of Linguistics,
University of Kansas
Lawrence, Kansas
KS-66044 3177
USA
rankin@ku.edu

Keren Rice

Department of Linguistics
University of Toronto
Ontario M5S 3H1
Canada
rice@chass.utoronto.ca

Eva Schultze-Berndt

Institut für Sprachwissenschaft
Karl-Franzens-Universität Graz
A-8010 Graz
Austria
schultze-berndt@ling.uni-graz.at

Dietmar Zaefferer

Institut für Theoretische
Linguistik
Ludwig-Maximilians-Universität
D-80799 München
Germany
zaefferer@lmu.de