elpublishing

Language Documentation
and Description

ISSN 1740-6234

This article appears in: *Language Documentation and Description, vol 1*. Editor: Peter K. Austin

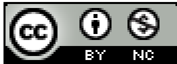
Genre, register and language documentation in literate and preliterate communities

WILLIAM A. FOLEY

Cite this article: William A. Foley (2003). Genre, register and language documentation in literate and preliterate communities. In Peter K. Austin (ed.) *Language Documentation and Description, vol 1*. London: SOAS. pp. 85-98

Link to this article: http://www.elpublishing.org/PID/009

This electronic version first published: July 2014



This article is published under a Creative Commons License CC-BY-NC (Attribution-NonCommercial). The licence permits users to use, reproduce, disseminate or display the article provided that the author is attributed as the original creator and that the reuse is restricted to non-commercial purposes i.e. research or educational use. See http://creativecommons.org/licenses/by-nc/4.0/

EL Publishing

For more EL Publishing articles and services:

Website:	http://www.elpublishing.org
Terms of use:	http://www.elpublishing.org/terms
Submissions:	http://www.elpublishing.org/submissions

Genre, register and language documentation in literate and preliterate communities

William A. Foley

One of the central questions posed by the papers in is volume is: 'how does language documentation differ from language description?' (Austin, this volume, Woodbury, this volume; Himmelmann 1998). Himmelmann (1998:166) draws the distinction as follows:

> "the aim of language documentation is to provide a comprehensive record of the linguistic practices characteristic of a given speech community. Linguistic practices and traditions are manifest in two ways: (1) the observable linguistic behavior, manifest in everyday interaction between members of the speech community, and (2) the native speakers' metalinguistic knowledge, manifest in their ability to provide interpretations and systemizations for linguistic units and events. This definition of the aim of a language documentation differs fundamentally from the aim of language descriptions: a language description aims at the record of A LANGUAGE, with 'language' being understood as a system of abstract elements, constructions, and rules that constitute the invariant underlying structure of the utterances observable in a speech community".

The main point of this paper is to take issue with such a rigid Saussurean dichotomy and to argue further that setting up the distinction in just this way is a reflection of linguistics' own, and largely unacknowledged (but see Harris 1980), Western ideology of language, as reflected in its cultural practices of normative dictionaries and reference grammars. i.e. a set of elements and a set of rules à la the Himmelmann (1998: 166) quote above. These cultural practices are largely, as Harris (1980) points out, the result of the rise of vernacular literacy movements in Western Europe in the sixteenth and seventeenth centuries, so that scientific thinking about language in linguistics is very much an artifact of its practitioners' own literate practices and the centuries long tradition of literacy which has preceded them. Literacy leads to the construction of powerful normative pressures in a number of areas: not only are prescriptive cultural products like grammars and dictionaries highly valued, but also the end products of literacy, literate texts, are given greater weighting over oral speech acts.

When linguists move to a fieldwork situation, they inevitably and mostly unconsciously brings their native ideology of language with them (to a significant extent canonized by their training in linguistics). But preliterate communities are unlikely to share their assumptions. Although recognizing the lack of literacy practices, many fieldworkers will nonetheless through the filter of their own native language ideology value some text types over others as sources of data, narratives over conversations, ritual language over

gossip, songs over curses, and use these as their primary bases for the construction of their own (ultimately normative, especially in the most endangered or moribund language communities) grammars and dictionaries. It is truly remarkable how most text collections in the languages of traditional communities around the world have such a high percentage of narratives, sometimes only these. Clearly, the linguists recording these have valorized them on the basis of their close analogy with literate texts in their own cultural tradition, an identification aided no doubt by cultural products like *Aesop's Fables* in that tradition. Whether the native community shares this differential evaluation is for the most part never discussed (or perhaps not even asked?). Woodbury (this volume) warns us about this trap in linguistic fieldwork and argues that the main goal should be the direct documentation of discourse, the full range of local genre types, and that language description is a secondary practice. I applaud this, but go further: language description is not only a secondary, but a highly idealizing and **cultural** practice (in our sister disciple anthropology, there has been much discussion about this aspect of ethnography, its equivalent of language description; see Burridge (1973). Unfortunately the problem seems little recognized in linguistics, the result no doubt of this discipline's universalizing predisposition, but even the cultural basis of this universalizing predisposition needs querying). We must always be as mindful as possible of the largely unquestioned assumptions we bring to fieldwork out of our own normative literate tradition. Because of the highly idealizing nature of language description, the effect of our native language ideology on our products of description can be heavily disguised (even our glossing practices, the starting point of any analytic description, reflects ideology; see Becker (1993, 1995)). The only corrective I can suggest for our possibly misleading descriptive flights of fancy is the hermeneutic one (Gadamer 1975; Taylor 1985), a linguistic analog of 'thick description' (Geertz 1973): stay close to the full range of data, all register and genre types; avoid differential evaluation of some text types over others, but search out the native estimations and their rationales for such (and being mindful of the possible disguising, hegemonic effect of their own ideology of language in this regard; see Kuipers (1990)); and when developing a description on the basis of these data, be prepared for inconsistencies and contradictions. As the fieldworker understands the language practices of the community at a deeper and deeper level, they should gradually resolve some of these, but the language practices of any speech community, no matter how small or monolithic (I have worked with language communities with a grand total of 200 speakers!), are never homogeneous. Variation is the rule; don't paper over this.

Let me try to illustrate some of these points through my own personal experience. I have worked in New Guinea with a number of Papuan language speaking communities in the Sepik-Ramu lowlands region of the north coast for some 25 years. New Guinea is linguistically the most complex region on earth: in an area of some 900,000 km^2 (the size of the state of New South Wales) are spoken some 1200 languages or 20-25% of the world's total; a roughly equivalent amount of languages was spoken indigenously in the entire Western hemisphere, but in an area of 49 million km^2. Genetically diversity is also extremely high: although not definitively established, the number of distinct language

families cannot be less than 30. The population density is low (less than 5 million people in the whole region), and most languages are spoken by less than 3000 people, many by less than 1000, and a couple of hundred by less than 100. All communities were preliterate until less than a century ago, and none of the language communities with which I have worked have literacy skills in the vernacular language, although some younger people are literate in Tok Pisin, and a few have rudimentary literacy skills in English that they have picked up in English medum schooling. Consequently, all native language texts that I have collected in these communities have been oral. This is in marked contrast to another language community with which I have worked in recent years, the Tibetan refugee community in India and Nepal. Tibetan is, of course, a language with a millennium old literate tradition and one highly valued at that, with an enormous accumulation of literate texts over that period. Working with a literate language community in one sense makes the fieldworker's job easier; for a mere $10 one can go into a Tibetan bookshop in Kathmandu and buy a body of texts that would take a team of 10 fieldworkers 10 years of full-time work to record and transcribe in a preliterate community. But, of course, even in Tibetan, , given the principle of 'thick description' proposed above, such a body of literate texts would be a inadequate sample of language documentation for language description, regardless of perhaps our own (and in this case native Tibetan also) predisposition to valorize them. But in the specific Tibetan case, the centrality of the literate-oral division in language documentation is highlighted by its peculiar diglossic (Ferguson 1972) situation. The language of this body of valued literate texts would be Classical Tibetan, a language very different from the dialectal varieties of Spoken Tibetan, be they of those of Lhasa, Tsang, Ladakh, Kham or Amdo, some of which may be so divergent as to be distant languages (I am ignoring here another literate form of the language, Modern Literary Tibetan (Goldstein 1991), which has emerged in the last 50 years for use in everyday secular literate works like newspapers, and is much closer to Spoken Tibetan, particularly of Lhasa). Thus, in Tibetan, the literate-oral divide is indexed by language code, Classical versus a form of Spoken Tibetan. In fact in Tibetan, the peculiarities of its diglossia mean that linguists are not predisposed to write normative grammars of 'Tibetan', though some have tried (e.g. Hannah 1973). More typically, we find distinct grammars of Classical Tibetan (e.g. Bacot 1948; Beyer 1992; Hahn 1985) and Spoken Tibetan, usually the Lhasa dialect, but not always (e.g. Bell 1919; Denwood 2000; De Roerich 1958; Koshal and Misra 1979; Kretschmer 1995; Read 1934; Tournadre and Dorje 2003).

But in most literate communities the literate-oral divide is not so transparently indexed, the same 'language' being used, as in English. This has lead linguists to postulate a continuum between literate and oral texts types., and various linguists have probed the formal linguistic differences between the ends of this literate-oral continuum. Literate texts are claimed to be more detached and more grammatically integrated than oral ones, which are more involved and structurally fragmented. We have two parameters (Chafe 1982, 1985):

literate	**oral**
detached	involved
integrated	fragmented

Figure 1: Parametric Differences in Literate-oral Texts

Some linguistic reflexes of these parameters in texts are: for literate texts as detached, the frequency of passives and nominalizations, and for integrated, frequency of subordination and other hypotactic syntactic relations, marked information structure options like cleft sentences and high lexical density, i.e. a high number of participant NPs and PPs per clause, particularly the latter in English. For oral texts as involved, the relevant features include a high frequency of speech act centered, i.e. first and second person pronouns, tag questions and expressive and colorful vocabulary. Diagnostics of oral texts as fragmented include a high frequency of simple coordination with *and* and flow monitoring words like *well, anyhow,* etc., as well as a general preference for simple juxtaposition or paratactic syntactic relations for constituent linkage. Olson (1977, 1994) goes further and argues that literate texts, due to their decontextualized nature, are necessarily more explicit in terms of expressing their meanings in an unambiguous and 'autonomous' manner, a source no doubt of the features of textual integration mentioned above. In oral texts, background meaning can be gleaned from the surrounding context, but in literate texts, available meanings are much more beholden to what is written, and this calls for greater explicitness on the part of the author. Invoking Grice's (1975) Cooperative Principle, speakers can usually assume a much higher degree of cooperation from their interlocutors, due to various types of shared information, and hence a greater latitude in the employment of conversational implicatures, than can authors. As a consequence of assuming less cooperative readers, authors necessarily invoke implicatures less frequently and are thereby more explicit in their wording so as to make their meanings clear and 'autonomous'.

Ultimately Chafe's and Olson's views can be boiled down to a claim of greater metalinguistic awareness in literate language over oral language. Writers/readers are more reflexively aware of their language as they perform literate activities than speakers/hearers: the language of literate texts is more carefully and slowly framed or can be monitored more critically over time, practices, of course, which get codified in cultural products like grammars and dictionaries. But a question immediately poses itself: do we want to claim that these are solely the effect of literacy, to deny such metalinguistic awareness to speakers/hearers in preliterate communities? Cultural practices in language communities with which I have worked in New Guinea strongly suggest otherwise. Metalinguistic awareness in these preliterate communities seems to be fostered by a number of cultural practices, debating duels, name etymologies, ritual songs, clan oral histories, but let me just consider one here: producing a text under conditions of visual feedback, an analog, of course, of what one engages in in literate practices in a literate community, but not literacy

per se. Let me explain. Inspired by Berman and Slobin (1994), I undertook the same experiment in the Watam speech community, a language of some 750 speakers spoken in two villages between the mouths of the Sepik and Ramu rivers on the north coast of New Guinea. I presented the children's book *Frog, Where are You?* (Mayer n.d.) to a number of Watam speakers and recorded their responses. *Frog, Where are You?* is an unusual children's book in that the story is completely narrated through a sequence of pictures; there is no language text, however basic, accompanying the illustrations. The book was given to Watam speakers to work through, and then they were asked to tell the story, again if they liked, working through page by page, an option all consultants took. Hence, the actual illustration was in front of the narrator as he narrated the events depicted on that page, but the overall text has the sequential event structure of a typical Watam narrative. The version of the text to be discussed here was given by a man in his late 30s, a first language speaker of Watam, but bilingual in Tok Pisin. He is not literate in Watam (no members of the language community are), although he has basic literacy skills in Tok Pisin due to six years of village based schooling in English (he has no real literacy in English, although the effects of schooling itself (and certain types of vernacular literacy practices) on metalinguistic awareness has been demonstrated in Scribner and Cole (1981). The language of his prompted narrative differs, as we shall see, in a number of important respects from a traditional Watam narrative, even those from the same speaker and other educated male speakers of the same age cohort, and these differences parallel those of the literate-oral continuum discussed above. This prompted narrative will be contrasted with a traditional Watam narrative about the origin of the moon. The narrator was a woman in her 50s, again a first language speaker of Watam and a second language speaker of Tok Pisin, although less proficient in that language than the first narrator. She has no literacy skills in any language and no formal education and overall could be regarded as a more 'traditional' member of the language community than the first narrator.

I start with the opening sentences of the two texts. The opening sentences of a narrative are typically information rich: setting the opening scene and introducing participants. Explicitness would seem to be called for, as seemingly little can be taken as given here at the beginning. Look first at the opening of the prompted narrative:

(1)

oraN	an	namtiN	i	ma	na	kiau	nambe	ep-tan
bush	this	child	a	3SG	POSS	dog	with	stay-PRES

'a boy is living with his dog in the bush'

miNga	na	endau	nakan	endau	na	nik	kaind
3DL	POSS	house	big	house	POSS	inside	thing.PL

ero-ero	wak-tan
various	be.at-PRES

'various things are inside the big house of those two'

The first sentence introduces the overall setting of the story, the bush, and its two main participants, a boy and his dog. The sentence contains 3 NPs to provide all this information in a rigid verb final structure, a high lexical density. The second sentence is similar: the specific sub-setting of the first scenes, the house, is introduced, as well as its contents, a crucial one of which will be the next major participant to be introduced on the following page, the frog. Again the sentence is a single verb final clause, and one with high lexical density. There are 3 NP constituents, the first two of which are in apposition to each other: 'their big house, inside of (the) house'. Note the high explicitness of integrated syntactic structure of these two sentences, a structure very close in fact, given its typological differences, to the English translations.

Now consider the opening sentences of the traditional narrative (R = realis; IRR = irrealis; RED indicates reduplication):

(2)

min	manda	Narap	an	wak	Nga-yak-na
3PL	today	tomorrow	this	sago	FOC-pound-IRR

Nga-mo-r	saN-mitena
FOC-do-R	go-HABIT

'everyday they want to go to work sago'

kamot	mbwoini	ni	miNga	saNga-r	yakaka
couple	two	DL	3DL	go-PAST	pound (RED)

'a married couple went and worked (sago)'

miNga	yak-mbe	iNba	Ngune	min	iri	mamai-na-nde,
3DL	pound-DEP.IRR	late	very	3PL	come	finish-IRR-really

miNga	iNba,	ero	an	Nga-momo
3DL	late	custom	this	FOC-do (RED)

'those two will work and very late, they (the others) will have already come, they (will come) late; they do thus'

These exemplify a much more typical Watam sentence structure than those of (1). Although it initiates the entire narration, the first sentence does not introduce any identifiable participants, just *min* 3PL 'they' or in this case perhaps a better translation would be an impersonal 'one', i.e. 'one does' or 'villagers do'. The temporal setting is introduced by 'today', 'tomorrow', meaning 'regularly, habitually', also identified by the tense suffix *-mitena* HABIT on the final main verb. Although final, the verb is actually a typical Watam serial verb construction, 'do/feel/become' 'go', meaning 'want to go', which in turn governs a purpose clause marked by the irrealis, 'to pound (work) sago'. The sentence actually contains two clauses, and each clause only contains one information bearing NP, *manda Narap* today tomorrow 'regularly' in the main clause and *wak* 'sago' in the purpose clause, in marked contrast to the up to 3 NPs in the clauses in (1).

Only in the second sentence are the main protagonists introduced, a married couple. The fact that it is a couple, i.e. two people, is heavily emphasized, by including the numeral ‘two’ with the word *kamot*, which means ‘married couple’, a seemingly redundant option, but also by the addition of the dual number marker *ni* in the NP and the use of the dual pronoun *miNga* in apposition, although this appositional use of a number marking pronoun with an NP is quite common in Watam. The importance of the restriction to ‘two’ is clearly to set this couple in contrast to the *min* ‘they/villagers’ of the first sentence; the reason for this will become apparent in the next sentence. Finally, the sentence again ends in a serial verb construction, ‘go’ ‘pound’, with the verb root for ‘pound’ *yak-,* reduplicated to indicate repeated action. Note that it is not irrealis as in the first sentence; this indicates a contrast between ‘wanting to go to work sago, but maybe not doing so’ versus ‘going and working sago’.

Having set the scene and the contrast between ‘they’ and ‘those two’, the couple, the next sentence introduces the story line in a complex, paratactic structure typical of traditional Watam oral narratives. Up to the first comma, which is used to transcribe a distinct falling pitch and short pause, there is a clause chain structure. Clause chaining is a structure in which separate clauses are strung out, one after another, but are dependent on the final clause for certain inflectional features, for instance tense, as in this example. The verbs in the clauses preceding the final clause have reduced possibilities of inflection, sometimes just a DEpendent marker.; only the verb of the final clause carries the full range of inflections typical of an independent sentence. Because this sentence is irrealis, *-na* IRR on the final verb, the marker of the verb of the dependent clause is *-mbe* DEP.IRR (DEP.R will be *-r*, as in following examples, although I will gloss it just as R there). Note the parallelism and contrast set up between ‘they two’, the couple, *miNga* 3DL, and ‘they’ *min* 3PL: ‘they both pound (sago) and very late, they (everyone else) will have really already come [‘come’ ‘finish’, another serial verb construction], they both (will come) late; (they both) repeatedly do this custom/behavior’. Note how weakly integrated, indeed fragmented, the structure of this sentence is, typical of oral language according to Chafe (1982, 1985). Note also the missing constituents, verb and subject NP, in the last two clauses. Both of these features are quite different from the information bearing, NP rich and tightly structured sentences of (1) and might suggest (1) was from a literate text, but remember both of these texts were delivered in an oral channel.

Now let me move on to the body of the texts and consider material drawn from there. (3) presents two sentences from the middle section of the prompted narrative:

(3) (a)

aes	an	tok	bibrak	mbo	Nga-birka-r-a
father	this	bum	sit.NOM	OBL	FOC-sit-R-DEP

itiN	ma	na	kiau	kiau	an	un	nakan	an	mbo
son	3SG	POSS	dog	dog	this	jar	big	this	OBL

kukurtamak	nakae	ndo-r	ma-iri-tak

frog	toward	see-PAST	PROG-go.down-PROG

‘the father (of the dog (i.e. boy)) sat down on a chair, while the son (i.e. dog) looked down on the frog in the jar’

(b)

pusirka-r-a,	irki-r-i,	kiau	na-(a)n	mo	moko-r-o,
jump-R-DEP	go.down-R-DEP	dog	POSS-this	DAT	hold-R-DEP

kiau	an	ma	me	ndika	yamsi-r-i,
dog	this	3SG	DAT	a.little	lick-R-DEP

ma	me	ndo-r-o	ma	me	digi-r	‘u	yaoN	ki?’
3SG	DAT	see-R-DEP	3SG	DAT	ask-PAST	2SG	good	Q

‘he followed him down and cuddled the dog; the dog licked him a little and he looked him over and asked him: ‘you ok?’

These sentences are longer and more complex than those from the opening, a typical feature of Watam texts. Consider (3a), another clause chaining structure, this time in realis with *-r* R on the dependent verbs and the final verb *ndo-r* ‘see’ inflected as PAST. There are two clauses, the first ending in the dependent verb inflected with *-r* R and the DEPendent echo vowel, which repeats the final vowel of the verb stem, in this case *-a* (the stem is *mbirak-* ‘sit’). Both clauses are information rich, having high lexical density with many NPs and PPs; the second has no less than 5! The boy and his dog are set up in a kinship relationship of ‘father’ and ‘son’, quite commonly extended like this is Watam culture. The first clause is relatively simple: ‘the boy sat on a chair’ (a chair is a ‘bum-sitter’). The second is very dense, with multiple appositions: ‘the son (i.e. dog), his dog, this dog…’. The clause ends in a serial verb construction, ‘see’ ‘go down’, ‘look down’, with the PROGessive aspect (*ma- ... -tak)* realized on the second verb in the construction, but the overall PAST tense on the first. (3a) is a highly atypical Watam sentence, far too much lexical density compressed into too few clauses, although this is a typical feature of literate language (Halliday 1985).

Now consider (3b). This scores much better on the question of lexical density per clause. The first line is a good exemplar: 3 chained clauses with only a single NP among them, and the first two clauses have no overt NPs or PPs at all, again quite typical. The subject ,‘boy’, of each, ‘jumping’ and ‘going down’ (interestingly not a serial verb construction, but expressed unusually in two chained clauses) is completely recoverable from the context of the previous sentences. The missing subject in the next clause, ‘(the boy) cuddles the dog’ is also fully recoverable, as it is the same as that of the previous clauses. However, in the next clause in the chain (all verbs in these dependent clauses are marked by *-r* R plus the DEPendent echo vowel), the subject shifts to the dog, which is overtly mentioned, and the boy as object appears as the pronoun *ma* 3SG. An alternative would be to elide the object NP, but this runs counter to the strong tendency in this text to explicitness, which also showed up in the lexical density of (3a). In the next two clauses

the subject reverts to the boy without any overt mention nor is the object, the dog, unambiguously expressed, only via the pronominal *ma* 3SG. In principle, the subject and object are not being unambiguously expressed, and this could be proposed as a contradiction to the claimed explicitness of this text. But look closer: the last verb is a speech act verb, *digi-* 'ask', with a direct speech act as a complement. Only humans speak (at least in this text!), so the boy is only possible subject here, and the object pronoun *ma* 3SG must refer to the dog. And the preceding clause clearly leads up to this question: '(he) looked him over and asked him (about his health)'. There is no real vagueness here; to specify 'boy' as subject is unnecessary and would strike Watam speakers as quite awkward stylistically. But what is really remarkable about (3b) is the string of 6 chained clauses in such an otherwise short sentence (contrast (4b) below) and the lack of **any** serial verb constructions. This is a very marked option and indicates a choice for explicitness and a transparent coordinate syntactic structure over compactness and paratactic linkage. Also each of these chained clauses is rigidly verb final.

Finally, consider these two excerpts from the middle of the traditional narrative:

(4) (a)

karir	namtiN	an	irki-r,	irki-r	kor
then	child	this	go.down-PAST	go.down-PAST	canoe

'then the child went down, went down to the canoe'

yokpaka-r	irki-r,	anup	aNgi-r	ausu-r	saNga-r-a,
push-R	go.down-R	oar	take-R	row-R	go-R

simuk	mbo	mbirka-r	ma-ndo-tak
point	OBL	sit-PAST	PROG-see-PROG

'(he) shoved (it, the canoe) down (into the water), took (his) oar and rowed away and sat down on the point and watched'

(b)

miNga	irki-r	wak	an	awe-r	mamai-r	irki-r-i,
3DL	go.down-R	sago	this	wash-R	finish-R	go.down-R-DEP

ndoNar	an	mo	insaNga-r	kor-paNan	Ngune	Ng(a)-atki-r,
moon	this	DAT	bring-R	canoe-head	very	FOC-put-R

gektaN	an	kondkai-r,	kor-paNan	an	aNgi-r
brightness	this	exceed-R	canoe-head	this	put-R

miNga	an	gektaN	an	mbo	ausu-r	ma-iri-tak
3DL	this	brightness	this	OBL	row-PAST	PROG-come-PROG

'they both went down, finished washing the sago and went down, bringing the moon, putting (it) right on the prow of the canoe; it was very bright and lit up the front of the canoe and they rowed and came by its brightness'

The first short sentence of (4a) introduces the main character of this scene, a boy. He goes down and then the going down is repeated with the goal, a canoe. Note that the goal NP

kor 'canoe' follows the verb (this is not an afterthought construction because no pause separates the verb and the NP, which would necessarily be the case in such a construction). Watam is actually not rigidly verb final at all; although the verb is most commonly last, a wide variety of NP and XP role types can follow it, including object NPs, but one would never discover this by studying the text prompted by *Frog, Where are You?* Now look at the next sentence. What is really remarkable in comparison to (3) is the high preponderance of serial verb constructions, no less than 3 in 3 chained clauses (the first not marked by the DEPendent echo vowel, which is not required, but the type of falling pitch transcribed by the comma is sufficient) and the lack of over NP participants. The number of events narrated is about the same as that of (3b), but instead of 6 chained clauses and no serial verb constructions, there are 3 chained clauses and 3 serial verb constructions. Further, the serial verb construction of the second clause, 'take' 'row' 'go', 'take and row away' is a more complex string than the simple 'see' 'go down', 'look down', of (3). What is salient here is a preference for compression and parataxis over explicitness and overt coordination. Also, the lexical density of (4a) is very low. The sentence contains 7 verb roots and only 2 NPs, while (3a) contains 3 verb roots and 7 NPs - an almost opposite ratio. NPs with discoursally established referents, whether subject, object or oblique NPs, are completely elided (e.g. first clause). NPs of new information are overtly mentioned (e.g. second and third clauses), but only one per clause, following a clear discourse preference rule: only one XP per clause and that bearing new information (Du Bois 1987).

Example (4b) illustrates this principle well and is a fine exemplar of a typical sentence in a traditional Watam oral narrative. Again, in contrast to sentences from the prompted narrative, the ratio of verbs to NP and PPs is high, 11 verbs to 8 NPs/PPs. Complex serial verb constructions are used, e.g. 'wash' 'finish' 'go down', 'finish washing and go down'. There are 7 chained clauses, and all but the last have at most a single XP. The last has two because it exhibits a change of subject from the previous clause, although the subject is only referenced by the pronoun *miNga* 3PL. Because the language lacks switch reference morphology to signal changes of subject across chained clauses, it is commonly signaled by overt mention of the subject NP, but this is not required. If the subject holds constant across chained clauses, as in the beginning clauses of this sentence, then overt mention of the subject is proscribed, although not impossible.

In summary, although Watam has no literate tradition and these texts were delivered in a completely oral channel, they contrast in a number of features typical of texts drawn from the ends of the literate-oral continuum. The text prompted by *Frog, Where are You?* is lexically dense, explicit and syntactically integrated, with relatively fixed word order and neat, set coordinated clause chains, all salient features of literate texts. The traditional oral narrative, on the other hand, is low in lexical density (lexical elaboration, if anywhere, is in the complex serial verb constructions), highly implicit, with much information needed to be recovered from context, and often structurally fragmented and freer, with more variable word order and extensive use of parataxis or juxtaposition, as in the serial verb constructions. These are all distinctive features of oral texts. Both texts

are oral in an absolute sense, but the prompted text had the visual stimulus of the book and its illustrations in front of the narrator, and this undoubtedly lead to a greater reflexive awareness of the participants and events of the story as he narrated it and yielded the structural patterns exemplified (the effects of his 6 years of schooling cannot be ruled out as a contributing factor here, although other texts from the same consultant exhibited typical oral narrative structures quite different from this, as did texts from other educated male speakers of the same age cohort).

There are a number of important implications from this short study, but I will mention just two. First, what are the data that we employ for our language description? Note that if we used data like those derived from our prompted text to describe Watam, we would advance a number of false generalizations: (1) the language is rigidly verb final; (2) the language only has simple, rather lexicalized, binary serial verb constructions; (3) object NPs are rarely elided; (4) clause chaining is favored as linkage device over verb serialization; and (5) clauses are NP rather verb heavy. Clearly, if we are to produce rich 'thick descriptions' of a language, we need to have a very extensive data base, drawn from a wide range of genre types. We need a rich corpus to produce even adequate descriptions. All descriptions, grammars and dictionaries, are inherently normative, whether we as responsible linguists wish them to be or not. And to the extent that our corpus is limited, we straitjacket the knowledge of the language we pass on to future generations, no small failing for the many isolated, politically weak and endangered language communities our work is commonly undertaken with. Our work will be normative for these communities. Unfortunately, what we describe of the threatened languages of many small language communities is often all that will be passed on of them to their descendents. We owe it to them to provide the 'thickest description' we are capable of. But even in literate industrialized language communities, this effect should not be underestimated. The work of linguistics typically valorizes the literate registers of the language of middle class speakers: take a look at the examples and grammaticality judgments of any linguistics journal. However much we might disown it, linguists work, typically unwittingly, to preserve the normative hegemony of this kind of language. What grammar of English is really adequate for the structures of ordinary colloquial conversational English? And which linguists have more academic prestige: theoretical linguists who concoct abstruse examples only possible (if that) in the highest registers of literate language of upper middle class speakers or sociolinguists who study casual conversation? The whole of linguistics is normative, even in our own unendangered language communities; how much greater a risk does our role carry in the poor, threatened, endangered language communities of the developing world.

Second, there are important implications in educational work. In the last five years or so, the Department of Education of Papua New Guinea has moved to promulgate a policy of vernacular medium education in the first few years of schooling rather than the previous English only policy. This is an important step, greatly to be applauded, and one that hopefully in the long run will aid the long term survival of the many small endangered languages of that nation. But I sound a word of caution. I wonder whether the use of

Watam in the village school (if it comes to this; it hasn't yet) and Watam literacy will promote over time a valorization of structures like those of the text prompted by *Frog, Where are You?,* rather than those of traditional oral narratives. Given the similarity of this text to that of literate texts elsewhere in the world, I suspect it might. And ultimately this is likely to lead to the devaluation of the rich and creative poetic structure of traditional oral narratives and perhaps their loss. This would be a great tragedy, a loss of the variety of human creative expression, not only of the Watam language community, but the world. The only way to counterbalance this, I think, is to be aware of the potent normative effects that linguistic work and tradition and the schooling practices derived from this have upon us and our understanding of the world. Our linguistic habitus (Bourdieu 1990, 1991) is powerfully compelling. I remember a review by John Haiman (1986) of Jeff Heath's (1984) excellent grammar of Nunggubuyu, in which he commented on a text about the frilled lizard included as an appendix to the volume. He said that couldn't imagine why any speech community would find this narrative interesting and commented how much more linguists needed to know about language in order to imagine this. We owe the communities we work with an obligation to expand our imaginations.

References

Austin, P. this volume. Introduction: some issues for discussion.

Bacot, J. 1948. *Grammaire du tibétain littéraire*. Paris: Adrien Maisonneuve.

Becker, A. 1993. The elusive figures of Burmese grammar. In Foley, W., ed, *The Role of Theory in Language Description,* 29-46. Berlin: Mouton-de Gruyter.

Becker, A. 1995. *Beyond Translation: Essays toward a Modern Philology.* Ann Arbor, MI: University of Michigan Press.

Bell, C. 1919. *Grammar of Colloquial Tibetan.* Calcutta: Bengal Secretariat Book Depot.

Berman, R. and Slobin, R. 1994. *Relating Events in Narrative: A Cross-linguistic Developmental Study.* Hillsdale, NJ: Lawrence Erlbaum.

Beyer, S. 1992. *The Classical Tibetan Language.* Albany, NY: State University of New York Press.

Bourdieu, P. 1990. *The Logic of Practice.* Stanford, CA: Stanford University Press.

Bourdieu, P. 1991. *Language and Symbolic Power.* Cambridge, MA: Harvard University Press.

Burridge, K. 1973. *Encountering Aborigines.* Oxford: Pergamon Press.

Chafe, W. 1982. Integration and involvement in speaking, writing and oral literature. In

Tannen, D., ed, *Spoken and Written Language: Exploring Orality and Literacy,* 35-53. Norwood, NJ: Ablex.

Chafe, W. 1985. Linguistic differences produced by differences between speaking and writing. In Olson, D., Torrance, N., and Hildyard, A., eds, *Literacy, Language and Learning: The Nature and Consequences of Reading and Writing,* 195-123. Cambridge: Cambridge University Press.

Denwood, P. 2000. *Tibetan.* London: School of Oriental and African Studies.

De Roerich, G. 1958. *Parler de l'Amdo.* Rome: Serie Orientale.

Du Bois, J. 1987. The discourse basis of ergativity. *Language* 63.905-855.

Ferguson, C. 1972. Diglossia. In Giglioli, P. *Language and Social Context,* 232-251. Harmondsworth: Penguin.

Gadamer, H. 1975. *Truth and Method.* Boston: Seabury Press.

Geertz, C. 1973. *The Interpretation of Cultures.* New York: Basic Books.

Goldstein, M. 1991. *Essentials of Modern Literary Tibetan.* Berkeley, CA: University of California Press.

Grice, P. 1975. Logic and conversation. In Cole, P. and Morgan, J., eds, *Syntax and Semantics* 3.41-58. New York: Academic Press.

Hahn, M. 1985. *Lehrbuch der klassischen tibetischen Schriftsprache.* Bonn: Indica et Tibetica Verlag.

Haiman, J. 1986. Review article of Jeffrey Heath, *Nunggubuyu Myths and Ethnographic Texts, Nunggubuyu Dictionary and Functional Grammar of Nunggubuyu.* Language 62.654-663.

Halliday, M. 1985. *Spoken and Written Language.* Waurn Ponds, VIC: Deakin University Press.

Hannah, H. 1973. *A Grammar of the Tibetan Language, Literary and Colloquial.* Delhi: Cosmo Publications.

Harris, R. 1980. *The Language-Makers.* London: Duckworth.

Heath, J. 1984. *Functional Grammar of Nunggubuyu.* Canberra: Australian Institute of Aboriginal Studies.

Himmelmann, N. 1998. Documentary and descriptive linguistics. *Linguistics* 36.161-195.

Koshal, M and Misra, B. 1979. *Ladakhi Grammar.* Neew Delhi: Motilal Banarsidass.

Kretschmer, M. 1995. *Erzählungen und Dialekt aus Südmustang.* (4 vols). Bonn: Vereinigung für Geschichtswissenschaft Hochasiens Wissenschaftsverlag.

Kuipers, J. 1990. *Power in Performance: The Creation of Textual Authority in Weweya Ritual Speech.* Philadelphia, PA: University of Pennsylvania Press.

Mayer, M. n.d. *Frog, Where are You?* New York: Dial for Young Readers Publications.

Olson, D. 1977. From utterance to text: the bias of language in speech and writing. *Harvard Educational Review* 47.257-281.

Olson, D. 1994. *The World on Paper: The Conceptual and Cognitive Implications of Writing and Reading.* Cambridge: Cambridge University Press.

Read, A. 1934. *Balti Grammar.* London: Royal Asiatic Society.

Scribner, S. and Cole, M. 1981. *The Psychology of Literacy.* Cambridge, MA: Harvard University Press.

Taylor, C. 1985. *Philosophy and the Human Sciences.* Cambridge: Cambridge University Press.

Tournadre, N. and Dorje, S. 2003. *Manuel de Tibétain Standard.* Paris: Langue et Civilisation L'Asiathèque.

Woodbury, A. this volume. Defining documentary linguistics.