Language endangerment, language documentation and capacity building: challenges from New Guinea

WILLIAM A. FOLEY


Link to this article: http://www.elpublishing.org/PID/018

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
Language endangerment, language documentation and capacity building: challenges from New Guinea

William A. Foley

1. Introduction

The New Guinea region, with its extreme linguistic diversity, over a thousand languages in a land area of 900,000 km$^2$, presents the urgent task of language documentation with an especially demanding challenge. Not only is the sheer number of languages daunting, but the actual number of genetic linguistic stocks is very high, at least thirty, but probably more, and the typological profiles of the languages are extremely varied (see Foley (2000) for an overview). While language endangerment is not as advanced as in some other areas of the world such as Australia and the Americas, most of the languages of New Guinea are spoken by very small speech communities, under a couple of thousand speakers and in many cases much less, and the inexorable advance of economic development, modernization and globalization threaten their viability as elsewhere, so that the current linguistic richness will certainly be greatly depleted by the turn of the next century.

The linguistic complexity of the region is most clearly exemplified along the north coast of New Guinea. With about a quarter of the land area, fully half of the languages are found there, and most of the language families. Particularly in the west of Sandaun Province and along the adjoining border with the Indonesian province of Papua are found many small language families and half a dozen or so language isolates, all with very small numbers of speakers. This is undoubtedly a residual zone, where small speech communities speaking ancestral languages have clung on in somewhat marginal lands, while more aggressive peoples speaking languages of larger expanding language families have surrounded them and surely assimilated many.

In this complex, shifting linguistic situation, language endangerment is unlikely to be a monolithic process, and it is therefore important to survey some of the paths different speech communities have taken. Consider the case of Karawa, a language spoken in one village by a total of 63 speakers (Ferree 2000). To its west lies the closely related Bouye language, spoken in six villages by eight hundred speakers. The speakers of Karawa are giving up their language in favour of their numerically superior close relative, as children are growing up only able to speak Bouye. This development was no doubt favoured by the fact that Bouye has traditionally been the contact language in intervillage communication between the sole Karawa speaking village and the Bouye speaking ones.

For a more complex development, consider Kopar village. The traditional language of this village was the Kopar language, spoken in it and in two other villages for a total population of fewer than three hundred. However, this language is gravely endangered; there are no fluent speakers under 45, and there has been no transmission now for nearly two generations. All villagers between 20 and 45 speak Tok Pisin as their first language. Starting from around 1991, all Kopar children have been educated in the local community school in neighbouring Watam village. The language of this village is Watam, very distantly related to Kopar (on a par with English and Hindi), and spoken as a first language by all Watam villagers. Through their immersion in a Watam speaking community, Kopar children have acquired Watam natively and have imported it into Kopar village. Though time will tell how stable this development is, Kopar village seems well on its way to a shift from a Kopar speaking community to a Watam speaking one, but through an intermediate generation that spoke neither language natively.

The loss of a vernacular language to a national language like Tok Pisin or Indonesian is, not unexpectedly, the most common cause of language endangerment in New Guinea. But the pathways that lead to this are not everywhere the same; each case requires a thorough linguistic and ethnographic study of the forces at work in the ongoing process of language replacement. Kulick (1992) presents such a thorough ethnographic description of the cultural beliefs behind a language shift from the vernacular language to Tok Pisin in Gapun village in the swampy area near the mouth of the Sepik River. The total population of the village is under a hundred. The language is an isolate with no known relatives, and is not spoken elsewhere. All adjoining languages are typologically very different from the Gapun vernacular. Like Kopar village, there are no younger fluent speakers; Tok Pisin is their sole native language. Gapun vernacular is viewed negatively, associated with the pre-contact traditional culture, a way of life stigmatized as backward by the villagers. Tok Pisin, on the other hand, is associated with wealth, modernization, the Papua New Guinea nation-state and its political-legal structures and the wider world, all positively valorised. There are other cultural ideological factors at work here too. The vernacular is linked to wilfulness and self-centeredness, psychological forces said to be holding back the modernization of the village, and, through these, to femaleness, for women are believed in the local ideology as more strongly exemplifying these traits than men. Conversely, Tok Pisin is paired with the knowledge and social responsibility necessary for the successful modernization of the village, and, as such, is tied to maleness, again reflecting cultural ideology about the socially cohesive roles of men. This complex amalgam of cultural beliefs has led to the very rapid transition from Gapun vernacular to Tok Pisin in a little over a single generation. Interestingly, somewhat similar processes are at work in Yimas village, but here have not resulted in the same outcome, at least not yet. What seems to be emerging in Yimas village is a gender-based diglossia. Younger women speak Yimas vernacular and Tok Pisin, but younger men use Tok Pisin. This is because girls at a very
early age are taken daily on all-female morning fishing expeditions, and Yimas vernacular is the language of choice among all-female groups. Consequently, girls grow up knowing the vernacular. But boys stay in the village from a comparably early age, playing in all-male groups, where Tok Pisin is the favoured code, and they thereby fail to learn Yimas vernacular. Whether this pattern is stable over the long term remains to be seen.

The factors which determine whether a speech community maintains its traditional language or shifts to another are complex, as the above case studies demonstrate, and to a certain extent remain mysterious. Consider the contrasting cases of Watam and Awar villages. Both are villages along the north coast of New Guinea, separated by about 60 km, and both came under European influence about a century ago, during the German colonial period. Their languages are closely related, on a par with Romance languages. Yet, the linguistic situations are quite different. In Watam village there is stable bilingualism; both Watam vernacular and Tok Pisin are spoken, and children are acquiring Watam vernacular as their first language. In Awar village children are largely not learning the vernacular, but are growing up monolingual Tok Pisin speakers. Why two villages in the same ecological and cultural zone, with parallel post-contact histories and closely related languages, have such different sociolinguistic profiles is a puzzle. Clearly, we need many more careful ethnographically sophisticated studies of speech communities in transition to understand the hows and whys of language endangerment.

Furthermore, the processes of language shift are not necessarily a sudden abrupt change. Even speech communities in which the vernacular seems healthy, with cross-generational language transmission like Watam, may exhibit telltale signs of language stress well before true language replacement sets in. This is exemplified in the language spoken by younger Watam speakers. The conservative Watam of older speakers exhibits very complex allomorphy for the pluralization of nouns, the vestiges of an earlier noun class system (Foley, in press); this is simplified and regularized by younger speakers. For instance, the plural of *rua* ‘coconut’ is *ruigg*, but younger speakers commonly use *ruaygar*, adding the general unmarked plural suffix to the singular form. The categories of Tok Pisin also affect the Watam of younger speakers. Conservative Watam has two verb roots, *kari-* ‘diswant’ and *yase-* ‘tired’. The meanings of both of these are expressed in a single lexeme *les* in Tok Pisin. Younger speakers increasingly use *yase-* to cover both meanings on the model of Tok Pisin, ousting *kari-* as a verb root in the process. In the neighbouring distantly related language Murik, these processes of language simplification and regularization across generations has been so thorough-going that the language, as currently spoken, is almost unrecognizable from that found in a grammar of the language written by a missionary merely a hundred years ago (Schmidt 1953).

These are all changes in language form, typically the core focus of interest of linguists’ fieldwork. Other changes may be less obvious at first, but are just as drastic in
the overall ecology of the speech community. The Watam language contains hundreds of lexemes for kinds of plants and dozens for types of birds, but few younger speakers know more than a handful of these. Subtle semantic contrasts in verbal collocations are being lost. And many younger speakers know little of or have much interest in the vast body of traditional legends, stories and other types of indigenous verbal art. Indeed, in the Yuat River language Biwat almost the entire corpus of traditional oral literature has not been passed on. It is as if the New Guinea equivalent of Grimm’s Fairy Tales or Aesop’s Fables or perhaps even the Icelandic Sagas has been lost without record, and this loss is a tragic ongoing process throughout New Guinea speech communities.

2. The task of documentation

While it is true that the situation of language shift may not be as dire in New Guinea as some other places in the world, it is still very serious, and the diversity and complexity of the region presents us with some very major problems in language documentation. The task of documenting New Guinea languages while they are still spoken in vibrant speech communities, with their full range of lexical knowledge and grammatical structures and rich array of genre types, is urgent. Few professional linguists work in the region; few languages are well described, and none fully documented, all this in an area where fully 20-25% of the world’s languages are spoken. We are confronted with a situation not unlike that facing those scholars who documented North American languages at the turn of the last century, and we can profit from a look at some of their research protocols. The founder of the Americanist tradition, Franz Boas, set up a framework for the documentation of Amerindian languages that was faithfully followed by his students and still stands as a model for language documentation even today. This consists of a detailed grammar and a voluminous dictionary plus, most importantly, a large corpus of transcribed and annotated texts. Because of the technological limitations of their day, Boas and his students collected texts by relatively slow dictation from native speakers. This ruled out spontaneous texts and multi-party impromptu conversations, so that the texts collected and published were typically of more formal genres like narratives, songs, prayers, recipes, etc. With the blessings of our modern technology, we are not limited to these, nor should we be; the recording and analysis of spontaneous speech events and conversations is vital to full language documentation (Himmelmann 1998), but it is still noteworthy how narrative-heavy most text collections of endangered languages remain even today.

But Boas presents a model for language documentation in still another – and perhaps even more valuable – way, namely through his long-term and close collaboration with native speakers. Although a staggeringly prolific publisher, Boas’ greatest contribution to anthropology and linguistics is undoubtedly his documentation of the culture of the Kwakiatl Indians and their language Kwak’wala. All in all, he published over 4,000
printed pages of primary texts in the Kwak’wala language, in a wide array of genres, a job of language documentation perhaps unmatched for any other language of a traditional or tribal community, anywhere in the world (Boas 1910, 1921, 1925, 1930, 1935; Boas and Hunt 1902, 1906). But, importantly, it wasn’t Boas who collected and transcribed these texts in the vast majority of cases, but George Hunt, a native speaker of Kwak’wala who Boas trained to write the language phonetically. Boas paid Hunt and encouraged him to collect as much primary material as possible, and it is this enormously productive collaboration that we have to thank for the unexcelled corpus of Kwak’wala texts available to us.

The Boas-Hunt collaboration model is one we would be well advised to adopt for the daunting task of language documentation in New Guinea. This task is impossible, simply too vast, for the small number of professional linguists to accomplish without these productive collaborations. We need to train, encourage and support native-speaker collaborators as co-researchers. But we need to improve on Boas’ training of Hunt and particularly in one crucial area of methodology. Boas and Hunt’s texts were written in the pre-phonemic period, and this, unfortunately, does somewhat compromise this accessibility. Current speakers of Kwak’wala report difficulties in comprehension, in large part due to an awkward overdifferentiated orthography for the vowels. Kwak’wala is a language with six phonemic vowels, a standard five vowel system /i e a o u/ plus schwa /ə/ (Grubb 1977), but Boas wrote and taught Hunt to write no less than 17 distinct vowels, basically the full range of allophonic variants. As with Northwest Coast languages generally (Bagemihl 1991; Kinkade 1997), the status of schwa is particularly problematic; it has a very wide range of allophones and determining whether it is underlying or epenthetic is difficult, features shared surprisingly with many New Guinea languages. Boas (1930: xi) in fact complains about defects in Hunt’s writing of the vowels, but this reflects often Boas’ misunderstanding of Hunt’s attempts at a phonemic writing of the vowels. Consider, for instance, Hunt’s writing of yəxs because-3SG in contrast to Boas’ yIkx (Boas 1930: xiv). The vowel is phonemically schwa, either underlying, or more likely, epenthetic (e.g. /yxs/), but Boas writes it with its allophone [i] which occurs in the environment of palatals. Hunt’s native intelligence led him to hit upon the notion of the phoneme as a psychological reality (Sapir 1949) in spite of Boas’ tutoring, and this notion of distinctive contrasts, a difference which makes a difference (Bateson 1972) or the emic-etic contrast in the sense of Pike (1967), is the single most useful idea we can teach our native speaker collaborators as they collect materials for a fuller documentation of their languages. But we can also learn a great deal from the way in which they document their languages, as Hunt’s insightful use of schwa demonstrated above.

Let me illustrate with two examples from my own fieldwork in the Sepik region of New Guinea. We too can gain valuable insights from native speakers’ attempts at writing
their language. I have mostly worked with speakers who have no literacy skills in any language, but in most villages there has always been a small handful of speakers who have at least rudimentary literacy skills in Tok Pisin. Asking these speakers to write down vernacular items using the spelling conventions of Tok Pisin has often proved valuable. I once asked a well educated speaker of Iatmul, who spoke both Tok Pisin and English and had a high degree of literacy skills how he would write the Iatmul word for ‘tree’, phonetically [mi]. I fully expected him to say mi, like the Tok Pisin word for ‘I’; in fact, I even prompted him as such. To my considerable surprise, he said “no, that’s wrong, I think it should be mwy”. This reflects directly the generally accepted underlying Iatmul representation /m+wi/, as the high, front and back and mid, front and back vowels in this language are the surface realizations of underlying central vowels plus the corresponding semivowel, e.g /i+ y/ → [i], /a+w/ → [u], /a+y/ → [e], /a+w/ → [o] (Foley 1986). The w in his spelling was his attempt to render /i/, for which there is no symbol in Tok Pisin orthography.

Now let us look at a more extended example. These are snippets from a classic Yimas text, in fact, the creation legend, known almost by heart by all mature villagers, which I recorded a number of times. I asked a well educated native speaker who had lived a long time in Port Moresby and spoke both Tok Pisin and English and had good literacy skills in both to listen to one of these recorded versions and to write what was said verbatim. Table 1 presents 9 (discontinuous) lines from this running text (see below for explanation).

1  T  mirim  tat  kambunakuan  mirim
    F  mrm  tat  kampunawkwan  mrm
    m-rm  tat  kampunawkwan  m-rm
    that-DL  PN  PN  that-DL
   ‘Those two, Tat and Kampunawkwan’

2  T  tatukuan  krana  tauna  tut
    F  tantukwan  krana(na)taunantut
       alone  3PC.S-DUR-sit-DUR-REM PAST
   ‘They always lived alone’
3 T tai min tampinumun taibi na ilbunkarak
F tai mn tampqumun taimpi nailpuŋkarak
tai m-n tampqumun tai-mpi na-il-nurj-kara-k
then that-SG PN see-DEP 3SG.S-down-walk-IRR
‘Then he saw (them) at Tampqumun and walked down’

4 T kikay minta krana wapalki taiak
F kikay mnta kranawapalmpitaik
kikay mnta kra-n(n)a-wapal-mpi-tai-k
PN then 3PC.O-3SG.A-DUR-climb-SEQ-see-IRR
‘Then Kikay climbed up and was watching them’

5 T kanta awgurumpumpin nanan apisambi ilimtut
F kanta aŋkurmpwimpn nanaŋapisampiirmtut
kanta aŋkurmpwimpn nanaŋ-apisa-mpi-irm-(n)tut
but PN 3SG.S-DUR-hang-SEG-stand-REM.PAST
‘But Aŋkurmpwimpn was hanging up’

6 T kran tai kirimpin pinan ikiakin
F krantaikmpn piaŋikŋkan
kra-n-tai-k-mp-n pia-n-i-kŋkan
3PC.O-3SG.A-see-IRR-time-OBL talk.0-3SG.A-tell-IRR-3PC.D
‘When he saw them, he said to them’

7 T “pangit kanta mintimpi pangra tauaringit?”
F “paŋkt kanta mntmpi paŋkrauwarŋkt?”
paŋkt kanta mnt-mpi paŋkra-tau-warŋkt
2 PC but that-ADV 2PC.S-sit-HAB-PC
‘But do you always live like that?’
Then this village fell out (of the sky) and lay (flat)

They continually shot (arrows) up; having continually shot, …

Table 1

T is the native speaker’s rendition; F is my phonemic rendition, using the orthographic conventions presented in the Yimas grammar (Foley 1991); and the next three lines present the phonemic transcription broken up morpheme-by-morpheme, an interlinear gloss and finally a free translation of each sentence. The phonemes of Yimas, while few, did present the native speaker with difficulties in three cases for which there are no symbols in the Tok Pisin orthography. Two of these are the palatal and velar nasals /ŋ/ and /ŋ/, rendered simply by n (line 3). The other is a pervasive epenthetic schwa, having a distribution not unlike that of schwa in Kwak’wala. Because this is predictable and epenthetic, my phonemically based orthography does not write it, but interestingly the native speaker did choose to do so, again not unlike George Hunt’s use of a symbol for Kwak’wala schwa, a practice probably favoured because in both languages schwa can be the only vowel in a word and as such carry primary stress. Look at the first word in line 1, phonetically [miɾim], which I write as mrm. The native speaker, however, writes schwa as i (the allophonic range of schwa includes [i] and [j], as well as schwa (see Foley 1991:44-45)). This results in orthographic under-differentiation, as both schwa and /i/ are written with same symbol; compare the first words in lines 1 and 4.

In other cases the native speaker’s writing overdifferentiates, but again in interesting ways. Yimas only has a single series of oral stops, usually voiceless, but voiced allophonically in certain environments, particularly after nasals, although even here a voiceless allophone is acceptable. I write all oral stops as voiceless, but the native speaker shows variation, reflecting the fact that he was well aware of a voiceless-voiced contrast, as exemplified in Tok Pisin orthography. Note that he often uses the symbols for the voiced stops after nasals (lines 1, 5, 7, 8), but he is by no means consistent, perhaps even more often using the symbols for the voiceless stops in this environment (lines 3, 5, 6, 7, 8,
9), a pattern of variation no doubt reflecting the allophonic status of this phonetic difference and its largely free variation pattern of distribution in this environment. He sometimes drops writing the nasal entirely, either with the symbol for the voiceless stop (lines 2, 4) or the voiced one (lines 3, 5). Note that these data are strong confirmatory evidence for the analysis which suggests that these nasal + stop phonetic sequences are in fact consonant clusters and not unitary prenasalized stop phonemes (Foley 1991:41).

There is a very interesting difference between the native speaker’s and my orthographic treatment of the word structure of complex verbs. This is illustrated in a number of lines, in which the native speaker has broken up verbs into a number of orthographic words, whereas I have written them as one (lines 2, 3, 4, 5, 6, 7, 8, 9). There are two main factors which account for this difference. The native speaker typically writes the complex of unstressed pronominal prefixes and aspectual prefixes which precede the verb root(s) as separate words (lines 2, 3, 4, 5, 6, 7, 9). This is surprising because this sequence of prefixes is not a word on either morphological or phonological grounds: they are necessarily bound, e.g. cannot stand on their own, and they cannot carry stress, as all phonological words must. However, this practice is exactly parallel to a usage of George Hunt commented on by Boas (1930: xiv). Kwak’wala has a set of pronominal agreement/case markers which behave as suffixes or enclitics, as evidenced by the phonological processes they undergo as a result of the preceding forms. Hunt wrote these as separate words, exactly as our native speaker does here. What is behind this parallel practice for treating apparently bound pronominal forms by two native speaker writers of otherwise typologically very different languages is a mystery, but clearly there is something which needs to be understood here.

The other major difference between the native speaker’s and my orthographic conventions for verbs as words concerns serial verb constructions (lines 4, 5, and 8). I write serial verb constructions as single words, for on morphological grounds they clearly are: e.g. pronominal agreement affixes for core arguments occur as both prefixes and suffixes, flanking the entire serial verb chain; negation shifts erstwhile pronominal agreement prefixes to suffixes following the complex of serialized verb stems, etc (see Foley 1991:80-87). The native speaker, on the other hand, seems more sensitive to phonological criteria. He writes each verb stem in a serial verb construction as a separate word, reflecting the fact that each carries its own primary stress. This is a strong diagnostic for wordhood for the class of nouns, where both morphological and phonological criteria coincide: a noun carries a single primary stress and takes only suffixal inflection as a whole. But it is not diagnostic for verbs because the morphological and phonological criteria conflict: complex verbal words like serial verb constructions take inflection, both prefixal and suffixal, as a whole, but they carry multiple primary stresses; or verbs take a single primary stress, but the affixal clusters are cut loose as free floating forms. One orthographic solution to this dilemma is not more correct than the other. Mine clearly
reflects my linguistic training and the weight that training has led me to give to notions of the contrast and distribution of forms. The native speaker’s solution is probably a textbook example of the kinds of effects of metalinguistic awareness highlighted by Silverstein (1981). Because nouns in Yimas are the only major part of speech that can stand on its own without inflection or derivation, they are highly accessible by Silverstein’s (1981) metrics to conscious metalinguistic awareness, and the cluster of formal properties exhibited by nouns become paradigmatic, in this case for orthographic wordhood.

The lesson of the collaboration of Boas and Hunt and the instructive meditations on the structure of Iatmul and Yimas by native speakers are vital to our project of language documentation. The active involvement of native speakers as co-researchers in language documentation projects is essential not only because this vast task needs their engagement, but also because they have a lot to teach us about their languages.

3. References


