

Ministry of Scientific Research and Innovation

WEH ORTHOGRAPHY GUIDE

by

Phil Davison

SIL

B.P. 1299, Yaounde, Cameroon

November 2009

Abbreviations

sg	Singular
pl	Plural
imp	Imperative
n	Noun
adj.	Adjective

Acknowledgments and Preamble

When compared to the work done on Aghem (where an SIL researcher spent at least 18 months in the Aghem area, and during that time worked with more than 100 Aghem speakers, many of them teachers, to develop an orthography guide), this current work on Weh seems quite inadequate. For one thing, the work has taken place almost exclusively in Yaounde: for another, these orthography suggestions are based on a phonology which dealt principally with the pronunciation of words in isolation. In addition, we are very much aware that the work has largely been done with the help of just one or two language resource persons. Having said that, it is impossible to undervalue the importance of the contributions of a number of people, foremost among whom is Ferdinand Asei. His vision for the development of the Weh language, and his enthusiasm and self-sacrifice in providing countless examples, are the main reason that this work was possible at all. Special thanks must also go to Mei Eugene Mou for the encouraging interest he has shown in his language, and for the real impetus he gave to the work, particularly during an intensive week in Bamenda (June 2009). Last but not least, we must recognise the contributions of Simon Buh and Chrysantus Kum, who were especially helpful in getting the language research started.

While acknowledging the limitations of what we have been able to do, it is our fervent hope that the Weh community will benefit from the great amount of work previously done by their Aghem ‘cousins’. It cannot be denied that there are differences between the two languages, not least in their phonology (and phonology plays a significant – but not exclusive – part in determining the orthography of a language). But there are definite areas where the suggestions made in the Aghem orthography guide should continue to be of great help to those working on Weh, particularly concerning word division.

One final word: it is with some ‘fear and trembling’ that this current work is presented. If the authors of the Aghem guide, after years of work by many people, saw fit to refer to their orthography proposals as ‘provisional’, how much more should we. I am more and more convinced that orthography development cannot – and should not – be undertaken essentially by a single person, and certainly not an ‘outsider’ to the language community. In addition, an orthography has to be far more than simply a reflection of the phonology of the language. In this regard, the current work is certainly deficient. As the proof of a pudding is in the eating, so the true test of any orthography is in its acceptance by the language community and its use over time. Some of the decisions elaborated herein, those informed by the General Alphabet of Cameroon Languages, are likely to remain valid: others, with time and greater wisdom, may well be changed. Be that as it may, these suggestions are humbly offered in the hope that they will prove valuable to the Weh community in the ongoing development of the Weh language.

TABLE OF CONTENTS

Abbreviations	ii
Acknowledgments and Preamble	ii
1 INTRODUCTION	1
2 ALPHABET	2
2.1 Consonants.....	3
2.1.1 Nasal-Consonant complexes	4
2.1.2 More double consonants.....	7
2.1.3 Consonants plus ‘w’	7
2.1.4 Other consonant combinations.....	8
2.1.5 Consonant orthography rules	8
2.2 Vowels	9
2.2.1 Vowel length.....	10
2.2.2 Vowel diphthongs	10
2.2.3 Vowel orthography rules	10
3 TONE.....	11
3.1 Lexical tone contrasts marked by diacritics	11
3.2 Grammatical tone contrasts	13
4 WORD DIVISION.....	14
4.1 Associative noun phrases.....	15
4.2 Compound words	18
5 PUNCTUATION AND CAPITALISATION	20
6 NEED FOR FURTHER TESTING AND RESEARCH.....	20
References.....	21

WEH ORTHOGRAPHY GUIDE

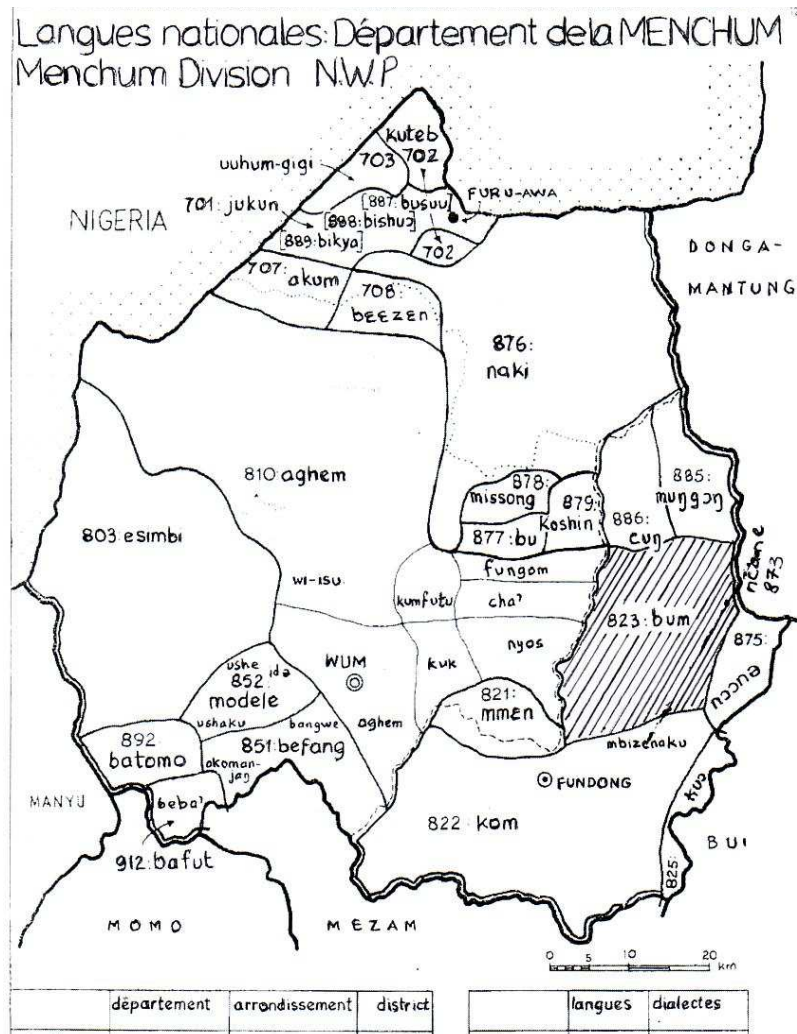
by Phil Davison

1 Introduction

This paper provides suggestions towards an orthography for Weh, a Grassfields Bantu language spoken by about 7,000 people in the Menchum Division of the North West Province, Republic of Cameroon. The SIL Ethnologue classifies Weh as: Niger-Congo, Atlantic-Congo, Volta-Congo, Benue-Congo, Bantoid, Southern, Wide Grassfields, Narrow Grassfields, Ring, West, and there is an additional comment: ‘similar to Aghem’. Both Weh and Aghem are listed in ALCAM under ‘langues de la zone 8’ under the number 810 (Aghem). The Weh language appears under the double heading ‘wi-isu’, which is subdivided as follows: “wi (= Weh < adm.) and isu (= esu, umusu)” (p.69). As indicated in the ALCAM entry, Weh is the name given by colonial administrators to the language, and incidentally to the main town of the Weh-speaking area. For the native speakers, the language is [kíwí], the [ki-] being a noun class prefix (class 7).

Whereas Aghem has received a good deal of attention (most notably from Larry Hyman in ‘Aghem Grammatical Structure’, 1979), the only work specifically on Weh known to us is an unpublished manuscript by Roland Kiessling entitled ‘A Proposal for a Practical Orthography of the Weh Language’, dated May 2002. The suggestions it contains are quite different from those given in this present document, most significantly because Kiessling avoids using any character that cannot be typed on a standard keyboard. The result is something that is very far removed from the guidelines given in the General Alphabet of Cameroon Languages. While making no claim to perfection, we can however state two things: the suggestions in this document are based on a significant phonological study of the Weh language, and they conform to the way in which the General Alphabet is currently used.

The following map showing the languages of Menchum Division (including wi-isu) is taken from ALCAM (section B, p.137):



2 Alphabet

The basic Weh segmental alphabet is made up of 28 consonants and 7 vowels. They are represented below in upper and lower cases:

A a, B b, Bv bv, Ch ch, D d, Dz dz, E e, F f, G g, Gb gb, Gh gh, I i, I i, J j, K k, Kp kp, L l, M m, N n, Ny ny, Ij ij, O o, O o, Pf pf, S s, Sh sh, T t, Ts ts, U u, V v, W w, Y y, Z z, Zh zh, '.

Provision should also be made for the exceptional letters **R r, P p** and **H h**, the first two for loan words which are in common usage in Weh (e.g. **trenjà** 'stranger, guest' from English; **põm** 'tap' from English 'pump') and the latter for the word **hayi** 'no' and various exclamations.

2.1 Consonants

The consonant phonemes, their allophones and graphemes are presented in the table below. We also show their use in various positions in the word.

Phoneme	Allophone	Grapheme	Initial position	Medial position	Final position
/t/	[t]	T t	tum news	nitɔŋ blow (horn)	—
/k/	[k]	K k	kab armpit	nikam squeeze	—
	[ʔ] ¹	’ ’	—	nidò’ò sit	finya’ eggplant
/b/	[b]	B b	bày father	ibi locust	—
	[p] ²		—	—	ughib bones
	[β] ²		—	nibiibi destroy	—
/d/	[d]	D d	dìn old	idîm burden	—
	[t] ³		—	—	kitàd spoon
/g/	[g]	G g	gòw shoe	logo cassava	—
/kp/	[kp]	Kp kp	kpèy namesake	tikpày balafon	—
	[kp] ~ [p ^h] ⁴		kpuuli dying	nikpu die	—
/gb/	[gb]	Gb gb	—	kigbòò match, flint	—
/pf/	[pf]	Pf pf	pfii chewing	nipfi chew	—
/bv/	[bv]	Bv bv	bvi dog	ibvîn corpse	—
/ts/	[ts]	Ts ts	tsi sun	titsɔŋ music	—
/dz/	[dz]	Dz dz	dzì voice	idzu mouth	—
/tʃ/	[tʃ]	Ch ch	chi open! (imp)	uchôd problem	—
/dʒ/	[dʒ]	J j	ji goat	ujim dream	—
/f/	[f]	F f	fib co-wife	kifâ valley	—
/s/	[s]	S s	sàŋ traditional shirt	nisàw stab	—
/ʃ/	[ʃ]	Sh sh	shi go out! (imp)	nishiili take away	—

Phoneme	Allophone	Grapheme	Initial position	Medial position	Final position
/v/	[v]	V v	—	uvid gun	—
/z/	[z]	Z z	ziŋ wind	izû quarrel	—
/ʒ/	[ʒ]	Zh zh	zhiimi singing	kizhi weeds	—
/ɣ/	[ɣ]	Gh gh	ghòw shell, skin	ighaŋ root, vein	—
/m/	[m]	M m	mwà arrow	nimày finish	tinyàm animals
/n/	[n]	N n	này mother	nitini accompany	kifin canoe
/ɲ/	[ɲ]	Ny ny	nyàm animal	fɲyi knife	—
/ŋ/	[ŋ]	Ŋ ŋ	ŋa'li pinching	nɲa' pinch	kinâŋ cocoyam
/l/	[l]	L l	logo cassava	kimali wedding	—
/w/	[w]	W w	waa child	kiwɔ hand	ighâw wing
/j/	[j] ⁵	Y y	—	nibayyi haggle	ikay charcoal

Notes:

¹ [k] occurs root-initially (first root consonant), [ʔ] medially or finally (i.e. second root consonant). Writing allophones with two different letters is not usually recommended, but this particular case is in common usage throughout Cameroon.

² [b] occurs root-initially and medially (where it is sometimes pronounced [β] when not in root-initial position): [p] (often inaudibly released) occurs utterance-finally.

³ [d] occurs root-initially and medially: [t] (often inaudibly released) occurs utterance-finally.

⁴ [p^h] occurs only before [u] (and may be pronounced [kp] in careful speech): [kp] occurs elsewhere.

⁵ [j] occurs medially only in its lengthened form.

2.1.1 Nasal-Consonant complexes

Before getting into the issue of nasal-consonant complexes in Weh (i.e. combinations of two different consonants), we need to look at a particular kind of consonant which is similar in some ways to those combinations. These are voiced consonants (stops or affricates, but not fricatives) which are preceded by a very short nasal element ('prenasalised'). Many consonants may be prenasalised, as in the following table:

Prenasalised consonant	Allophone	Grapheme	Initial position	Medial position	Final position
/ ^m b/	[mb]	Mb mb	mbà fence	timbò' shoulders	—

Prenasalised consonant	Allophone	Grapheme	Initial position	Medial position	Final position
/ ⁿ d/	[nd]	Nd nd	ndaw <i>house</i>	tindɔŋ <i>sweet potatoes</i>	—
/ ⁿ g/	[ŋg]	Ng ng	ngaw <i>antelope (sg)</i>	tingaw <i>antelope (pl)</i>	—
/ ⁿ mgb/	[ŋmgb]	Mgb mgb	mgbi <i>thorn</i>	tingbi <i>thorns</i>	—
/ ^m bv/	[mbv]	Mbv mbv	mbvin <i>bush</i>	timbvi <i>chickens</i>	—
/ ⁿ dz/	[ndz]	Ndz ndz	ndzɔŋ <i>moon</i>	indzi <i>cloth</i>	—
/ ⁿ dʒ/	[ndʒ]	Nj nj	nji <i>sheep (sg)</i>	tinji <i>sheep (pl)</i>	—

A very similar consonant complex, i.e. a nasal sound followed by another consonant, is created with one particular noun class prefix. The prefix in question (noun class 12) is a syllabic nasal which also assimilates to the point of articulation of the following consonant. (A list of all noun classes, with explanatory notes, can be found below ***) Being syllabic, this nasal element carries its own tone (usually high, as is the case for all noun class prefixes), and may be pronounced with a little more duration than the nasal element of the prenasalised consonants above. The following table shows examples of class 12 nouns with the syllabic nasal prefix:

Consonant complex	Allophone	Grapheme	Initial position	Medial position	Final position
/Nt/	[n:t]	Nt nt	ntay <i>gravel</i>	—	—
/Nk/	[ŋ:k]	Nk nk	nkiyi <i>beer</i>	—	—
/Nb/	[m:b]	Mb mb *	mbu' <i>chimpanzees</i>	—	—
/Nkp/	[ŋm:kp]	Mkp mkp	mkpid <i>pimple</i>	—	—
/Nts/	[n:ts]	Nts nts	ntsô' <i>salt</i>	—	—
/Ntʃ/	[n:tʃ]	Nch nch	nchɔŋji <i>groundnuts</i>	—	—
/Ndʒ/	[n:dʒ]	Nj nj *	njim <i>geckos</i>	—	—
/Nf/	[ŋ:f]	Mf mf	mfûŋ <i>flour</i>	—	—
/Ns/	[n:s]	Ns ns	nsa' <i>needles</i>	—	—
/Nz/	[n:z]	Nz nz	nzizi <i>praying mantises</i>	—	—
/Nɣ/	[ŋ:ɣ]	Ngh ngh	nghi <i>milk</i>	—	—

Consonant complex	Allophone	Grapheme	Initial position	Medial position	Final position
/Nl/	[n:l] ~ [l:]	Nl nl	nlĩ <i>drink</i>	—	—

* In the majority of cases in this chart, the syllabic nasal prefix occurs before consonants other than those which can be prenasalised. However, the syllabic nasal prefix *can* occur before /b/ and (rarely) /dʒ/: the examples given in the chart above are [mbúʔ] ‘chimpanzees’ and [ndʒĩm], ‘geckos’. This means that the written form for the nasal-consonant sequence will be the same for such words and those with prenasalised /^mb/ and /ⁿdʒ/: (i.e. **mb-** and **nj-**). There is a potential orthographic ambiguity here, but we believe that such cases will be rather rare, and any attempt to disambiguate them will lead to other complications.

NB: In the case of all of these syllabic nasals, the nasal element is not always present in all contexts: cf. [ńtsòʔ] ‘salt’ and [tsoʔ muŋ], ‘my salt’ because of a morphological rule that drops the prefix in certain environments. This is different from the prenasalised consonants given in the previous table: there the nasal part is always present, cf. [ndāw] ‘house’, [ndaw zuŋ] ‘my house’.

There is a further complication, in that there are class 12 nouns with root-initial prenasalised consonants or nasals. In contexts where the prefix (syllabic nasal) is present, we suggest doubling the nasal letter in the orthographic form, to represent both the prefix AND the following nasal element, as shown in the table immediately below:

Consonant complex	Allophone	Grapheme	Initial position	Medial position	Final position
/N ^m b/	[m:b]	Mmb mmb	mmbĩm <i>fleas</i>	—	—
/N ⁿ d/	[n:d]	Nnd nnd	nndàŋ <i>stools</i>	—	—
/N ^ŋ g/	[ŋ:g]	Nng nng	nngwḁ <i>palm rats</i>	—	—
/N ^m bv/	[m:bv]	Mmbv mmbv	mmbvi <i>oil</i>	—	—
/N ⁿ dz/	[n:dz]	Nndz nndz	nndzi <i>urine</i>	—	—
/N ⁿ dʒ/	[n:dʒ]	Nnj nnj	nnjim <i>bats</i>	—	—
/Nm/	[m:]	Mm mm	mmu <i>water</i>	—	—
/Nn/	[n:]	Nn nn	nnin <i>birds</i>	—	—
/Nj/	[j:]	Nny nny	nnyi <i>knives</i>	—	—

Note that if this doubling is not done, then there is no way of distinguishing certain pairs of words, such as ‘chicken’ and ‘oil’: they would both be written **mbvi**.

2.1.2 More double consonants

In addition to the double nasals mentioned above (which result from a syllabic nasal prefix before another nasal letter), there are other double consonants. These come about when a suffix is added to another word (typically verbs, but this can also occur with nouns). Thus, there is a rule that whenever two identical consonants occur in a single word in Weh, there is always a morpheme boundary between them. The effect on the pronunciation is a consonant which is noticeably longer than a single consonant, and because of this, the recommendation is to write such words with two identical consonant letters, as in the table below. Since these will never appear initially, their graphemes are not shown with capital letters.

Consonant complex	Allophone	Grapheme	Initial position	Medial position	Final position
/mm/	[m:]	mm	—	ninimmi <i>growl</i>	—
/nn/	[n:]	nn	—	nidinnì <i>pluck</i>	—
/ŋŋ/	[ŋ:]	ŋŋ	—	taŋŋi <i>report (n)</i>	—
/ll/	[l:]	ll	—	nizilli <i>listen</i>	—
/ww/	[w:]	ww	—	sìwwi <i>watery</i>	—
/jj/	[j:]	yy	—	nizàyyi <i>rest</i>	—

Once again, if this doubling is not done, then there is no way to distinguish word pairs such as **ninanji** ‘leave behind’ and **ninanŋji** ‘keep, store’.

2.1.3 Consonants plus ‘w’

Another kind of consonant complex is where the second element is the labial-velar approximant [w], which is always written in Weh as "w". Quite a few consonants may be followed by this sound, as in the following table:

Consonant complex	Allophone	Grapheme	Initial position	Medial position	Final position
/tw/	[tw]	Tw tw	twòd <i>blunt (adj.)</i>	—	—
/kw/	[kw]	Kw kw	kweyyi <i>changing</i>	fìkwò <i>belt, whip</i>	—
/bw/	[bw]	Bw bw	bwòlli <i>beating</i>	fìbwòd <i>cat</i>	—
/gw/	[gw]	Gw gw	gwavà <i>guava</i>	nìgwò <i>grind</i>	—
/ ⁿ gw/	[ŋgw]	Ngw ngw	ngwà’ <i>young man</i>	fìngwò <i>palm rat</i>	—
/fw/	[fw]	Fw fw	fwòlli <i>mixing</i>	kìfwà’ <i>work (n)</i>	—

Consonant complex	Allophone	Grapheme	Initial position	Medial position	Final position
/zw/	[zw]	Zw zw	zwòlì <i>itching</i>	izwôd <i>itch (n)</i>	—
/mw/	[mw]	Mw mw	mwà <i>arrow</i>	kimwa' <i>cockerel</i>	—

2.1.4 Other consonant combinations

There are a couple of other consonant complexes in Weh. Even though these are much less frequent than the case of consonants plus 'w' above, they are a part of the orthography. The letters for writing them are all contained in the alphabet, so these should not cause any problems for writers.

The first of these rarer consonant complexes is where the second element is the palatal approximant [j], which is always written in Weh as "y". Only a few consonants may be followed by this sound, as in the following table:

Consonant complex	Allophone	Grapheme	Initial position	Medial position	Final position
/tj/	[tj]	Ty ty	—	ityim <i>liver</i>	—
/kj/	[kj]	Ky ky	—	kikyì <i>baboon</i>	—
/lj/	[lj]	Ly ly	—	kilyim <i>comb (of cockerel)</i>	—

The second such complex is where the second element is the velar fricative [ɣ], which is always written as "gh" in Weh. Only a few consonants may be followed by this sound, as in the following table:

Consonant complex	Allophone	Grapheme	Initial position	Medial position	Final position
/bɣ/	[bɣ]	Bgh bgh	bghummi <i>hunting</i>	nibghum <i>hunt</i>	—
/fɣ/	[fɣ]	Fgh fgh	fghum tuɲ <i>my gossip</i>	tifghum <i>gossip</i>	—

These combinations of letters < bgh > and < fgh > are very unusual, but this is a reflection of the rarity of these sound combinations in Weh. If Weh readers and writers are unhappy about the suggestion made here, an alternative spelling could be adopted: **nibughum** 'hunt', **tifughûm** 'gossip'.

2.1.5 Consonant orthography rules

1. Do not write "p" or "t" at the end of words. If you hear these, write "b" and "d" respectively.
2. If you hear "p" before "u", write it as "kp".

3. A “p” on its own should only be written in loan words (such as **pɔm** ‘tap’, from English ‘pump’). Similarly, the letter “r” only occurs in loan words.
4. Do not write “k” at the end of words (always write ’ for this). Never write ’ initially.
5. If you hear a nasal consonant before “k” or “g”, always write “n”.
6. If you hear a nasal consonant before “kp” or “gb”, always write “m”.
7. Take care with initial nasals: there are times when a double nasal letter needs to be written if a syllabic nasal prefix is followed by a word which begins with a nasal letter.
8. If you hear a long (‘stretched’) consonant in the middle of a word, write two letters together: the consonants which can be long are “mm”, “nn”, “ŋŋ”, “ll”, “yy” and “ww”.
9. The only consonants that can appear at the end of words are ’, **b**, **d**, **m**, **n**, **ŋ**, **y** and **w**.

2.2 Vowels

The Weh language has 7 vowel phonemes, presented with their allophones and proposed graphemes in the table below. The sounds are shown as they occur in different positions of the word.

Phoneme	Allophone	Grapheme	Initial position	Medial position	Final position
/i/	[i]	I i	itam <i>fishhook</i>	iwiwi <i>mosquito</i>	ji <i>goat</i>
	[ɪ] ¹		—	kinyim <i>tongue</i>	—
/e/	[e] ²	E e	—	iwey <i>nose</i>	—
/i/	[i]	I i	—	kidɪŋ <i>fon</i>	ibvi <i>death</i>
	[ə] ³		—	kifɪn <i>boat</i>	itsɪŋi <i>truth</i>
/a/	[a]	A a	aghum <i>eggs</i>	nigham <i>bend down</i>	kidà <i>lamp</i>
/u/	[u]	U u	ukâ’ <i>mountain</i>	kituŋi <i>ear</i>	nikù <i>cook</i>
/o/	[o] ⁴	O o	—	kidzo’ <i>donkey</i>	logo <i>cassava</i>
/ɔ/	[ɔ]	ɔ ɔ	—	kisɔŋ <i>louse</i>	kiwɔ <i>hand</i>

Notes:

¹ /i/ has a centralised allophone [ɪ] which occurs in roots of CVC structure, particularly if the final C is [m]. Elsewhere it is realised as [i].

² Short /e/ only occurs followed by /y/, as in the example given here.

³ /i/ has two allophones, a more close [i] and a more open [ə]. The former occurs in roots with a CVC structure if the final C is /ŋ/, and in CV roots: [ə] occurs elsewhere.

⁴ /o/ occurs most frequently before glottal stop [ʔ], and only occurs in final position when the vowel in the first syllable is also **o** (e.g. **logo** cassava).

2.2.1 Vowel length

Weh has both short and long vowels, written as a sequence of two vowel letters. Long vowels are not as common as short ones, but examples for each one are given below (with the contrasting short vowels):

jii	<i>road</i>	ji	<i>goat</i>
aghee	<i>where</i>	aghey	<i>wives</i> ¹
idzfi	<i>termite</i>	idzi	<i>dirty</i> ²
kaa	<i>monkey</i>	ika	<i>egusi</i>
ukuu	<i>tail</i>	uku	<i>chairs</i>
kigbòò	<i>match</i>	boʔ	<i>bow</i> ³
indzòò	<i>intestine</i>	tò	<i>pot</i>

Notes:

¹ As noted above, short /e/ is always followed by /y/. Long /e:/ is pronounced [ɛ:].

² Long /i:/ is pronounced with the more open of the two central vowel allophones, i.e. [ɛ:].

³ As noted above, short /o/ occurs most frequently before [ʔ]. Long /o:/ is rare.

2.2.2 Vowel diphthongs

Diphthongs can be defined as combinations of two different vowels: in Weh, there are no such combinations in the orthography. As noted above, although it may *sound* as if two vowels come together, these are written as sequences of vowel plus either **w** or **y** (e.g. **ndaw** ‘house’, **ukây** ‘bed’).

The Aghem language has sounds which are known as ‘velarised diphthongs’ (Hyman pp.8-9), such as (Aghem) [bìyà] ‘two’. The situation in Weh is much more straightforward, and there is no reason to think of such sounds as diphthongs at all. They can be written quite simply as combinations of vowel and consonant letters, as in the chart below. It should be noted, however, that when the sound sequence [-iyə] follows either /k/ or /g/, a certain phonetic reduction takes place. The final two examples in the table below show both the phonetically reduced forms and the full orthographic forms.

Phonetic form	Orthographic form
[bìyà]	bighi <i>two</i>
[ítsíyó]	itsighi <i>fish</i>
[úkxə]	ukighi <i>money</i>
[ŋgyə]	ngighi <i>shelf</i>

2.2.3 Vowel orthography rules

Various rules concerning which vowels can occur in which positions are important to learn. Here is a partial list of vowel orthography rules:

1. The only vowels which can begin a word are “a”, “i” and “u”. These will in many cases be noun class prefixes: note that, as we have seen above, there are circumstances where

nouns appear without their prefixes. Note also that singular and plural forms of nouns have different noun class prefixes, and this is another reason why nouns have a tendency to look ‘different’ in different contexts.

2. The only time where two vowels can be written together is for a long vowel (i.e. two vowels of the same quality): two different vowels should not be written next to each other. If you hear “ai” or “ei”, write “ay” or “ey”. If you hear “au” or “ou”, write “aw” or “ow”.
3. An “e” should never be written on its own: it should be followed by “y”, or by another “e” (in which case it is a long vowel, as in **aghee** ‘where’).

3 Tone

Verbs in Weh show only a two-way contrast in citation form (between high and low). With nouns, we have identified six contrastive tonal melodies: three of them are pronounced on a level pitch: high [á], low [à] and something that is realised between the two. We will use the phonetic mid-level mark for this, i.e. [ã], but it should probably be thought of as a lowered high tone. In addition, there are two falling melodies, one which falls from high to low [â], and another which falls from high to lowered high [ã̂]. Finally there is a rising melody [ã̃]: this is not realised as low-high, but rather as low-lowered high. These melodies can all occur on noun roots of one or two syllables: in the latter case, a melody that changes pitch (falling or rising) will be ‘spread’ across the whole root (see the second of the tables below).

3.1 Lexical tone contrasts marked by diacritics

At this point in time, we have discovered numerous minimal tone pairs for words that contrast high and low (both nouns and verbs), and other pairs that contrast high and high-low. Some melodies are relatively rare (in particular the falling and rising melodies). Because minimal tone pairs are so rare between high and the fall high-lowered high, and there are none at all for the contrast high v lowered high, we propose an orthography where low is marked with a grave accent (e.g. à), the high-low fall is marked with the circumflex accent (e.g. â), and the others are left unmarked.

Tone	Phonetic	Grapheme (Diacritic)	Examples	Frequency
High	[á]	a	isi <i>eye</i>	e.g. very many nouns and verbs
Lowered High	[ã]	a	tĩmbvi <i>chickens</i>	e.g. some nouns
Low	[à]	à	kifi <i>lock</i>	e.g. many nouns and verbs
Low-Lowered High	[ã̂]	a	kĩtsow <i>grasshopper</i>	e.g. only a few nouns
High-Low	[â]	â	ifû' <i>basket</i>	e.g. many nouns
High-Lowered H	[ã̃]	a	ighañ <i>root, vein</i>	e.g. a few nouns

Note that in the table above (and the one below), the tone on the noun class prefix in each case is high: the contrastive tone indicated in the first column is found on the noun root.

The vast majority of noun class prefixes occur with high tone: in the cases where the noun class prefix tone is low, it can (and should) be marked. This will help to distinguish such pairs (or close pairs) as:

fĩmbu' [fĩmbũʔ] *banana* **ñtim** [ñtəm] *jiggers*
fĩmbù' [fĩmbùʔ] *tapping knife* **ntam** [ñtám] *fishhooks*

As mentioned above, these same melodies are also realised on nouns of different structures. In most cases the marking remains the same, but there are some minor differences. It will be helpful to see examples of longer words with these melodies:

Tone	Phonetic	Nouns with 2 syllables	Nouns with long vowels
High	[á(C)á]	ibɔŋɔ <i>reed</i>	waa <i>child</i>
Lowered High	—	—	—
Low	[à(C)à]	kisà'ì <i>comb</i>	kigbòò <i>match</i>
Low-Lowered High	[à(C)ā]	fĩmbù'u <i>bile</i>	indzòò <i>intestine</i>
High-Low	[á(C)à]	kighalà <i>weaver bird</i>	idzfi <i>termite</i>
High-Lowered H	[á(C)ā]	kituŋi <i>ear</i>	ukuu <i>tail</i>

Note that if a noun of two syllables or with a long vowel has a Low melody, then this should be indicated on both vowels: if the low mark is placed only on the first vowel, then there is no clear distinction between that and the marking for Low-High nouns. In this way, the Low-Lowered High melody (not indicated on a single syllable) is represented when it occurs on a longer root: the majority of Low-Lowered High nouns are of this latter type. Note also that the circumflex [â] is retained for those nouns with a High-Low falling melody on a long vowel: as shown in this table, the mark is placed over the first vowel. (We believe that this will be conceptually easier, for both writers and readers, than the alternative, i.e. having a low tone mark on the second vowel.)

There are many tonal changes that occur when words are put together in phrases. However, our proposal is to indicate the tone on all Weh nouns, verbs and adjectives as they occur in isolation (i.e. as they occur when pronounced by themselves), and it is these forms that should be written in the dictionary. Weh speakers should have no problem reading or pronouncing the actual phonetic pitch of entire sentences as long as they easily recognize the meanings of all the individual words that make up that sentence and any grammatical tone markings that might be required by the orthography (see following section).

It is fairly easy to find Weh words where the only difference between them is tone. For example, the word pairs below illustrate the most frequently contrasted tonal melodies:

ifow [ífw] *axe* **inyi** [íni] *knee*
ifôw [ífw] *leak* **inyî** [íni] *lake*

kítsa' [kítsáʔ] *mud*
kítsà' [kítsàʔ] *rattle*

nizu [nízū] *buy*
nizù [nízù] *fight*

kíkîm [kíkễm] *piece*
kíkîm [kíkễm] *log*

kíkù' [kíkùʔ] *small of back*
kíkù' [kíkùʔ] *deaf mute*

As mentioned above, there are rare tone pairs that contrast the less frequently occurring melodies. It is our feeling that these words do not need to be distinguished with diacritics and that they will be disambiguated by how they are used in context. In fact, the only orthographically ambiguous word pair that we have found so far is shown below:

ifow [íʔw] *axe*
ifow [íʔw] *leaf*

We do realise that our proposals mean that single-syllable nouns with FOUR melodies (High, Lowered High, Low-High and High-Lowered High) are all unmarked. However, as we have seen there seems to be very little risk of ambiguity in the orthography: for one thing, the rising and falling melodies are not all that common. It is our belief that 'less is better than more' when it comes to orthographical tone marks, and we have done our best to mark the essential distinctions. The whole area of tone marking is complicated, and it is quite likely that as the proposals outlined above are put into practice, Weh readers and writers will find there is a need for minor changes. This is something to be encouraged, but it should be done based on experience, in a careful and consistent manner, and with feedback from a qualified linguistics consultant.

An important decision regarding tone marking will have to be made in the future: how much of the vocabulary of Weh should be marked for tone? So far, we have referred only to nouns, verbs and adjectives needing tone marks (they being the only ones that participate in minimal tone pairs): this is one system that some languages adopt (and it means leaving unmarked such things as pronouns, grammatical particles, etc.). Another choice made for other languages is to mark ALL words of the language for tone (that is, they would be marked according to the system outlined above, either [à] if they have a low tone in citation form, or [â] if they have a high-low falling tone in citation form: otherwise they would be unmarked). Given that it may be very difficult to know exactly what the citation form of some particles is, it seems preferable to adopt the first of these options.

3.2 Grammatical tone contrasts

Many Grassfields Bantu languages use pitch differences not just for distinguishing between words (as above), but also for signalling specific grammatical meanings. This is an area which needs much more attention. It is hoped that this can be carried out as soon as possible, to help both readers and writers of the Weh language.

Given the closeness of Weh and Aghem, we expect that there will be similarities between the two languages in the area of 'grammatical tone'. A study of the Aghem orthography statement will be very useful (sections 3.1.2 & 3.1.3, looking specifically at the examples illustrating grammatical tone). It is not necessary to follow precisely what Aghem does to mark grammatical tone differences in the orthography: indeed, the suggestions we have made in the previous section are somewhat different from those contained in the Aghem orthography statement for the marking of lexical tone. But we do expect that the two languages might use grammatical tone in similar ways.

The kinds of phrases that can be distinguished in Aghem by tone alone are the following:

- 1) Present tense plus intentional aspect (e.g. ‘you want to eat’) v Future (later today) (e.g. ‘you will eat’ – later today).
- 2) Recent past (e.g. ‘he was sick’ – today) v Distant past (e.g. ‘he was sick’ – before today).
- 3) Indicative immediate past (e.g. ‘he ate fufu’) v Hortative (e.g. ‘he should eat fufu’).

(Linguistics consultants will be able to help with understanding the technical terms here and also with constructing comparable phrases in Weh.)

If the situation in Weh is found to be the same as in Aghem, there are various ways that the orthography can signal differences in grammatical tone. For example, Aghem uses a HIGH tone mark **á** for hortative phrases (the ‘acute accent’ is not used as a lexical tone mark in Aghem). Other languages adopt a mark which is completely distinct from accents, such as the dieresis (two dots over the vowel), e.g. **ä**. In Ngiemboon, hortative phrases are marked with exclamation marks both at the beginning and the end of the phrase (see Anderson’s ‘Précis d’orthographe pour la langue ngiemboon’): although this is unusual, it provides a mark right at the beginning of the phrase to help readers with the correct pronunciation of what follows. This latter suggestion could prove useful in Weh too.

As said above, we recognise the need for more work on the subject of grammatical tone. For the moment, Weh speakers should use the lexical tone-marking rules and make a careful note of any grammatical ambiguities that might need to be indicated in the orthography.

4 Word division

Like grammatical tone, this is an area which needs further study. Even at this early stage, though, there are some things which can be said: in general, those who will be making decisions about Weh word division would be wise to follow the principles already developed for Aghem (particularly section 4.1 of the Aghem orthography statement).

One specific thing we can mention is that, when present, all the normal Bantu noun class prefixes will be written as part of the nouns, as for the classes shown below:

Class	Prefix	Noun	Gloss
1	-	waa	<i>child</i>
2	a-	awaa	<i>children</i>
3	u-	ukây	<i>bed</i>
4	i-	ilîŋ	<i>bamboo</i>
5	i-	isi	<i>eye</i>
6	a-	asi	<i>eyes</i>
7	ki-	kiwɔ	<i>hand</i>
8	u-	uwɔ	<i>hands</i>
9	-	mbɔŋ	<i>cow</i>
10	ti-	timbɔŋ	<i>cows</i>

11	fi-	finin	bird
12	N-	nnin	birds

NB: The numbers for the classes in this chart do not conform exactly to the numbering used for all Bantu languages (although they are the same as those used for Aghem by Larry Hyman): our classes 10, 11 and 12 here are given the numbers 13, 19 and 6a by those who compare Bantu languages wherever they are spoken. We believe it will be much simpler to use the numbers 1-12 in the context of teaching Weh.

As previously mentioned (section 2.1.1), the prefix for class 12 is a syllabic nasal which assimilates to the point of articulation of the consonant which follows.

4.1 Associative noun phrases

There are cases when Weh noun roots appear without prefixes, e.g. **wə kuŋ**, ‘my hand’ (NB: ‘hand’ is **kiwə**). Another case in point is the first noun of an associative noun phrase, i.e. where two items are placed together in a relationship of ownership, e.g.:

fin ki finyim ‘the elder’s boat’ (literally: ‘the boat of the elder’)

Here, the first noun (the “possessed” item, or N1 for short) appears without the prefix that it would have if pronounced in isolation (**kifin**, ‘boat’). Note that it is followed by an associative marker **ki**: this marker resembles the prefix and always ‘agrees’ with the noun class of the noun it follows. Finally, the second noun (the “possessor” or N2) appears as it would in isolation, with its prefix. It is recommended to write each part of such phrases separately. It is very important to note that the tones of the two nouns in an associative noun phrase will always be written as they are for those same nouns when pronounced in their citation forms. Tones do have a tendency to change when words are placed in phrases, but in this particular case we believe it is very important to preserve the ‘constant word image’ for the specific nouns.

The example just given is relatively straightforward: this is however not always the case with associative noun phrases, particularly when an associative marker or the noun class prefix of N2 is a vowel (or both). Vowels have a tendency to merge into each other, and it can sometimes be difficult to ‘disentangle’ them, as shown in these two examples:

si i mbəŋ ‘the cow’s eye’ ([si: mbəŋ], from **isi** ‘eye’ & **mbəŋ** ‘cow’)

si a mbəŋ ‘the cow’s eyes’ ([sə: mbəŋ], from **asi** ‘eyes’ & **mbəŋ** ‘cow’)

Once again, we recommend writing each part of the phrase separately, even though this is not exactly how the phrase is pronounced in normal speech.

Our impression is that the vowel ‘disentangling’ is made easier if the writer bears in mind that there will, in the vast majority of cases, be *three* parts in an associative noun phrase (i.e. N1 without its prefix, then the associative marker which refers to N1, and then N2 with its prefix). Thus, when hearing [təwɔ:kət] ‘heads of slaves’, a useful procedure would be first of all to identify the two nouns in question, **utow** ‘heads’ and **ukɔd** ‘slaves’. Then it will be easy to understand that the long [u:] between the two noun roots is a running together of the N1 associative marker **u** and the N2 noun class prefix **u-**:

tow u ukôd ‘the slave’s heads’

It is difficult to give a complete list of associative markers at this time, as there are several complications that have not yet been fully explored. Since we expect there to be many correspondences between Weh and Aghem, perhaps the best we can do for now is show the associative markers of *Aghem*, and recommend more research to verify whether what is true for Aghem is also true for Weh. In order to form a complete picture for Weh, associative noun phrases should be constructed with nouns of each of the 12 classes in both ‘slots’ (N1 and N2): in addition, particular attention needs to be given to N1 nouns that end in vowels, since we know that vowels have a tendency to blur together.

AGHEM noun class	associative marker
1	u
2	a
3	u
4	i
5	i
6	a
7	ki
8	u
9	a
10	tî
11	fî
12	n

(NB: The noun class numbers in this chart and the one on the following page are those used by Larry Hyman in ‘Aghem Grammatical Structure’: as mentioned above, classes 10, 11 and 12 are often numbered 13, 19 and 6a by other Bantu linguists.)

Although we cannot as yet comment in any detail on the corresponding markers in Weh, there is one specific thing that can be said:

- We do note a difference with the associative marker for noun class 1. In Aghem, this marker is **u** (it will appear as such in the orthography between the N1 from class 1 and a following N2). However, this is not the case in Weh. The word for ‘child’ is **waa** (high tone), and the word for cow is **mbɔŋ** (lowered high): when they are placed together in an associative phrase, it might seem as if there is no marker present: we do however note a change of tone:

[wà: mbɔŋ] ‘calf’ (lit. ‘child of cow’)

Thus there is an audible marker (i.e. the change of tone on **waa** to LOW), but it is not easy to represent this in the orthography. (NB: *this shows that attention should also be paid to tone in any future research of the Weh associative noun phrase.*) Since it is important to preserve a ‘constant word image’ (to facilitate word recognition), we do not recommend indicating the audible low tone on **waa**, as it has a high tone in its citation form. There is no overt marker that can be written, and therefore this is the one case where an associative noun phrase will NOT have three specific parts. However, there is a way of indicating that the two nouns in this case are combined

in an associative phrase, by using the hyphen:

waa-mbɔŋ ‘calf’

This hyphen is recommended for use in the case of an associative phrase where N1 is from noun class 1: it can be seen as serving to indicate that there is an associative marker between the two nouns. In all other cases (it seems), there will be an overt marker present. (Note that this is a rather unusual usage of the hyphen when compared to other Grassfields Bantu languages: however, it does seem to be a useful device in this specific case in Weh.)

Another thing that adds to the complication is that some markers seem to have TWO forms, depending on the phonological context. Here are some examples:

- 1) [ŋgaŋ ziweɟ] ‘bridge of nose’, from **ingân** ‘edge’ and **iwey** ‘nose’. Here, the N1 noun has an **i-** prefix in citation form: it is interesting to note that the ‘concord consonant’ for noun classes with an **i-** prefix is **z-** (i.e. possessives, adjectives, demonstratives etc. following such nouns begin with **z-**).
- 2) [taj yitsi] ‘cooking stones’ (from **atay** ‘stones’ and **itsi** ‘fireplace’). In similar fashion, the concord consonant for noun classes with an **a-** prefix is **gh-**.
- 3) [naʔ mi ŋkala] ‘tomatoes’ (from **nnya** ‘egg-plants’ and **ŋkalà** ‘white man’). The concord consonant for noun class 12 is **m-**.

Once again, a comparison with Aghem is helpful. On page 25 of the Aghem orthography statement, the chart of associative markers shows how some of them vary. What is important is that these ‘alternative’ associative markers appear specifically before vowels and syllabic nasals. Here is a list of the Aghem associative markers in both forms:

AGHEM noun class	‘standard’ marker	marker pre V & N
1	u	u
2	a	ghi
3	u	wi
4	i	zi
5	i	zi
6	a	ghi
7	ki	ki
8	u	wi
9	a	a
10	ti	ti
11	fi	fi
12	n	mi

It will be vital to check to see if the Weh associative markers vary in similar ways. Going on the 3 examples given just above the table, it seems that they will be comparable, and if this is indeed the case, the written forms of the phrases in the above examples will be as follows:

- 1) **ngân zi iwey** ‘bridge of nose’
- 2) **tay ghi itsi** ‘cooking stones’
- 3) **nya’ mi ñkalà** ‘tomatoes’

One further aspect of associative noun phrases which may cause some difficulties concerns nouns which end in **-d**. As we have already seen (in the notes under the table of consonants), the citation form of such nouns may sound as if they had a final **-t**, but they should always be written with **-d**, e.g. [úvət] **uvid** ‘gun’. The added complication when such nouns form part of an associative noun phrase is that the final consonant sometimes sounds like **-l**:

[zəli finyim] ‘the elder’s name’ (from **izid** ‘name’ & **finyim** ‘elder’)

Our recommendation is that the final **-d** should be maintained in such cases:

zid i finyim ‘the elder’s name’

Keeping the final **-d** means that the noun root is always spelt in the same way (**zid**), whether the final sound happens to be [t], [d] or [l] in the pronunciation. This is another instance where the written form preserves a constant word image for readers, allowing them more easily to recognise the meaning of the word in its various contexts.

There is clearly more research to be done on this subject, and more decisions which will have to be taken on how to write certain things. The overall principle should be the one already adopted: to try as much as possible to preserve the form of the different parts of the phrase when they indicate the same meaning.

4.2 Compound words

In general, compound words in Weh are formed like associative noun phrases (i.e. N1 with no prefix + associative marker + N2 with prefix): in some cases it may be difficult to know if something is a compound or an associative phrase. It is however important to know which is which, because they should be written differently. (In this, we are following the good work that has already been done for similar cases in Aghem. For more details see the Aghem orthography statement, sections 4.1.6 and 4.2.2.)

As we have seen in the preceding section, associative noun phrases are to be written separately (usually in 3 parts). Compound words on the other hand should be written as one single word. What follows is drawn largely from the Aghem orthography statement (since their situation is the same as what we find in Weh, and they have developed some very helpful principles).

The word for ‘cowry shell’ in Weh is **siikighim** [si:kxim]. This comes from the words **isi** ‘eye’ and **ikighim** ‘crab’ (literally, cowry is ‘eye of crab’). In this case, we will *not* write this **si i ikighim** (as if it were an associative noun phrase). However, we *will* write **si i mbɔŋ** for ‘cow’s eye’, even though this is pronounced [si:mbɔŋ]. What are the reasons for these decisions?

Semantics (i.e. meaning) will provide one of the determining factors in deciding whether such constructions will be written as one word or as a phrase. In the case of ‘cowry shell’, there is clear consensus that the total signifies a single semantic idea or concept, and a native speaker does not derive the meaning from the sum of the parts. The second construction (‘cow’s eye’) on the other hand brings to mind three concepts: an eye, a cow, and a relationship of ownership or belonging. The following principles will be used to decide when to write single words and when to write noun phrases:

- 1) The meaning of the compound word should signify a single semantic concept (such as ‘cowry’).
- 2) The meaning of the compound word should not be immediately derivable from its parts. For example, the word **waa** ‘children’ can be coupled with the names of dozens of animals giving such results as **waa-mbvi** ‘chicks’, **waa-mbɔŋ** ‘calves’, etc. It would be a mistake to turn all of these very transparent constructions into separate lexical entries. On the other hand, although the word **mwa’kaazi** ‘butterfly’ is in some sense a combination of the words **kimwa** ‘cockerel’ and **kizi** ‘god’, a native Weh speaker isn’t thinking at all about cockerels or God when saying the word **mwa’kaazi**...
- 3) Although both compound words and associative constructions will permit the pluralisation of the first noun, only associative constructions may pluralise the second noun. For example **waa-mbvi** signifies the chick of one fowl while **waa-timbvi** immediately signals that the chick belongs to more than one fowl. In the case of compound nouns, attempts to pluralise the second noun will create either a non-existing word or a complete change in meaning. Taking the ‘cowry’ example from above, pluralising the second noun would result in **si i tikighim**, an associative construction that can only mean ‘eye of the crabs’. (The plural of cowry is actually **siikighim**, ‘cowries’.)

It was stated above that compound words usually have the same form as associative noun phrases: this means that the pronunciation in itself may not help to distinguish between them (as in the case of [si:kxim] ‘cowry’ and [si:mbɔŋ] ‘cow’s eye’). We have however discovered at least one case where there is a difference in the pronunciation. Consider the following two phrases:

[mwa?ka:zi] ‘butterfly’ (from **kimwa** ‘cockerel’ and **kizi** ‘god’)

[mwa? ki kizi] ‘God’s cockerel’ (from the same two nouns)

Both pronunciations are ‘good Weh’, but the meaning is different. The first, **mwa’kaazi**, is a compound (written as a single word), and the second, **mwa’ ki ki zi**, is an associative noun phrase (written in three parts). It is possible that there are other such helpful distinctions in the language that remain to be discovered.

Summary:

Determining whether a given construction is a compound or an associative phrase is one of the tasks of a language committee, as is developing a full lexicon (i.e. a list of Weh words). The work will not be easy and there will likely be some controversy since many constructions, unlike the examples above, will not fall so easily into one category or the

other. Those working on the lexicon are encouraged, when dealing with this particular topic, to ensure that their lexical decisions are soundly based on Weh semantics and not on the number of words found in the English translation. When the decision is taken to add a compound word to the lexicon, the committee should endeavour to have it reflect the surface form, with vowel-elisions, etc., rather than trying to reconstruct an underlying associative phrase. Due to the way vowels blend together differently in different environments, it is worth noting that some compound nouns will potentially have two quite distinct forms for singular and plural (e.g. **siikighim**, ‘cowry’, **siikighim**, ‘cowries’). The ‘irregular’ spellings should not be too critical since many languages possess a few words with different singular and plural forms (e.g. English ‘mouse/mice’, ‘tooth/teeth’).

5 Punctuation and capitalisation

The punctuation marks and the rules that govern the Weh language are the same as those for English. The punctuation marks used for Weh are therefore as follows:

1. Full stop (.)
2. Question mark (?)
3. Exclamation mark (!)
4. Comma (,)
5. Colon (:)
6. Semicolon (;)
7. Quotation marks (“ ... ”) and (‘ ... ’)
8. Parentheses (brackets)

As in English, capital letters should be used at the beginning of sentences. A capital is also appropriate for the first word of a direct quotation or after a colon (:), if what follows the colon is a complete sentence. Proper names, place names, titles and days of the week will also have capitals. In the case of language names, the Weh people might like to consider capitalising the root noun rather than the prefix, especially for their own language, i.e. **kiWi** ‘Weh language’. (Although this is not currently done in other Cameroonian languages, it is a common practice in other Bantu languages, e.g. kiSwahili.)

6 Need for further testing and research

We have made references throughout this guide to areas where more research is needed: it might be useful to summarise them at this point. Firstly, a decision will need to be made about how much of the Weh vocabulary should be marked for tone (just nouns, verbs and adjectives as we have suggested, or possibly every single word). An additional consideration in tone marking will be grammatical tone: we suspect that Weh makes some grammatical distinctions purely by tonal differences: if this is the case, the questions to be answered are; do those distinctions need to be indicated in the orthography, and if so, how. Thirdly, the whole area of the associative noun phrase is not completely understood (although neither, it has to be said, are many areas of the grammar). Of perhaps even more immediate importance than any of these is the fact that much of the non-basic vocabulary of the language remains to be investigated: a collection of around 800 basic words (containing mostly nouns and verbs) might be sufficient for a phonology write-up, but it certainly cannot meet the needs of any future language work.

That being said, it is our fervent hope that this proposal will be a first step in allowing Weh people to start reading and writing their language, that ways will be found to continue the research, and that improvements to these initial suggestions will be made as it is used.

References

- Anderson, Stephen C. 2007. Précis d'orthographe pour la langue ngiemboon. SIL Cameroon. Online version: <http://www.silcam.org/folder030401/page.php>
- Dieu, M. & P. Renaud, eds. 1983. Atlas Linguistique de l'Afrique Centrale (ALAC). Inventaire préliminaire : le Cameroun. Yaounde and Paris: ACCT/CERDOTOLA/DGRST.
- Hyman, Larry, ed. 1979. Aghem Grammatical Structure. SCOPIL 7. Los Angeles, University of Southern California.
- Kiessling, Roland. 2002. A Proposal for a Practical Orthography of the Weh Language. (Unpublished manuscript.)
- Lewis, M. Paul (ed.). 2009. Ethnologue: Languages of the World, Sixteenth edition. Dallas, Tex.: SIL International. Online version: <http://www.ethnologue.com/>
- Tadadjeu, Maurice and Etienne Sadembouo, eds. 1979. General alphabet of Cameroon languages. PROPELCA 1. Yaounde: University of Yaounde.
- Thormoset, David & Nelson Tshonghongi. 2003. Guide to the Aghem Orthography. Yaounde: SIL Cameroon.