

NOTES FROM THE FIELD

Minangali (Kalinga) Digital Wordlist: Presentation Form

Kenneth S. Olson

SIL International and University of North Dakota

Glenn Machlan

SIL International

Nelson Amangao

Christian and Missionary Alliance Churches of the Philippines

This paper presents a 207-item digital wordlist of Minangali, an Austronesian language spoken in the Philippines. The wordlist includes orthographic and broad phonetic transcriptions of each word, an English gloss, an individual WAV recording of each item, and metadata for resource discovery. An archival form of the wordlist was deposited into an institutional archive (the SIL Language and Culture Archives) and includes the original WAV digital recording, the original RTF wordlist form, descriptive markup encoding of the wordlist in XML employing Unicode transcription, and the metadata record. The presentation form was then generated directly from the archival form.

1. INTRODUCTION.¹ This paper presents a 207-item digital wordlist of core vocabulary in Minangali, a variety of Lower Tanudan Kalinga spoken in the town of Mangali, Kalinga Province, Philippines. Lower Tanudan Kalinga is an Austronesian language (Northern Philippine subgroup) spoken by approximately 11,000 people (ISO 639-3 code: kml, Gordon 2005). This presentation form was generated from an archival form of the data. The procedure we followed for creating both forms is detailed in Simons, Olson, and Frank 2007 and summarized below.

In addition to a description of the primary data in the form of phonetic transcription, we provide a documentation of the data in the form of digital audio recordings (cf. Himmelmann 1998), enabling the reader to verify and critique our transcription. This is important for this particular wordlist because Minangali has some unusual phonological phenomena, including the rare interdental approximant speech sound (to our knowledge, these are the first published recordings of the sound) and word-internal ...VCV... sequences in which the intervocalic consonant is the coda of a preceding syllable.

The materials included in this presentation of the data include the following:

- Wordlist: the present report, containing for each word an English gloss,

¹ We wish to thank Jan Allen, Fraser Bennett, Mike Cahill, Steve Parker, and two anonymous reviewers for helpful comments; Gary DuBois, Judy Kuntz, Steve Quakenbush, Dean Salter, Gary Simons, Translators Association of the Philippines, and SIL-Philippines for technical and logistic assistance; and SIL Asia Area for funding.

orthographic and broad phonetic transcriptions, and pertinent notes.

- Recordings: WAV² digital recordings of each item, accessible by clicking on the loud speaker icon next to the orthographic form of each word in the list below. Your web browser will attempt to play the recording with the sound program that is set up as the default WAV player on your system. On Windows computers, you can download the file by right-clicking on the icon and selecting “Save Target As...” from the pull-down menu.
- Metadata: a resource description of the data, useful for resource discovery.

The original wordlist materials included two items: an electronic wordlist form (in RTF format) and an 18.5-minute digital recording in WAV format. The wordlist form that we used presented 207 items, which is an amalgamation of the Swadesh 200 (Swadesh 1952:456–457) and Swadesh 100 (Swadesh 1955:124,133–137) wordlists, with some minor modifications (*thou* → *you sg.*, *ye* → *you pl.*, *person* → *man (human being)*, *woods* → *forest*, *berry* → *fruit*, *claw* → *fangernail*, and *right* → *correct*). For each item, the form provided a prompt in English and spaces for the transcription of the elicited form in both orthography and broad phonetic script. The third author translated the wordlist into Minangali, with some assistance from the second author.

The wordlist was recorded on April 17, 2006 in the recording studio at the SIL Center in Bagabag, Nueva Vizcaya Province, Philippines. During the recording session, the second author produced the English prompt, and the third author produced the target word twice. The recording was made using a Samson C01U USB studio condenser microphone connected to a notebook computer, using Speech Analyzer v. 2.7 for audio capture.

The third author is a sixty-year-old native speaker of Minangali. He has lived most of his life in Mangali, but he lived in Cebu for three years during his post-secondary education. He also speaks Ilocano, English, Cebuano, and some Tagalog.

The microphone and software allowed for recording at a sampling rate of 44.1 kHz and a quantization of 16-bit (i.e., standard audio CD quality). This is sufficient for technical purposes since it covers nearly all acoustic information pertinent to language, but it does not meet the generally-accepted recommended best practice of 96 kHz and 24-bit for archival data (Ladefoged 2003:18, 26; Simons et al. 2007:31; Plichta and Kornbluh 2002; IASA-TC03 2005:8). We recommend that field researchers record primary data at archival-quality rates if possible.

The draft transcription contained fields for five annotations: the item number, an English prompt, an orthographic transcription of the Minangali utterance, a broad phonetic transcription of the Minangali utterance using the International Phonetic Alphabet (IPA 1999), and additional notes. The draft transcription was revised by the authors, converted to a comma-delimited (CSV) file, and imported into TableTrans v. 1.2 software (Bird et al. 2002), where it was time-aligned to the original audio recording. This annotation was outputted to an XML annotation graph output and transformed into an XML descriptive wordlist format using an XSLT script.

The original electronic wordlist (in RTF format), the original WAV file, the XML

² Definitions of WAV, XML, XSLT, and other pertinent terms employed here are available at <http://emeld.org/school/glossary.html>.


descriptive wordlist, and a metadata record constitute the archival form of the wordlist (Machlan and Olson 2008). The metadata record follows the standard set up by the Open Language Archives Community (OLAC, <http://www.language-archives.org/OLAC/metadata.html>). A copy of the archival materials can be ordered on CD-ROM for a nominal fee from:

SIL Language and Culture Archives
7500 W Camp Wisdom Rd
Dallas, TX 75236 USA
archive_dallas@sil.org

















The presentation form of the wordlist was then generated from the archival form. An XSLT script was employed to convert the archival XML descriptive wordlist into an HTML presentation wordlist. Then, TableTrans was used to automatically create individual sound files corresponding to each of the segments identified in the transcription process for use in the presentation form.

The broad phonetic transcription requires a few remarks. First, two adjacent identical consonants represent a long consonant. We write them as two consonants in order to allow for stress to be marked properly on forms such as item 33: *abobba* [abob'ba] 'short'. Second, two adjacent identical vowels represent two distinct vowels occurring in separate syllables, such as in item 20: *poos* ['poos] 'few'. Third, some words have an unusual pattern in which there is a syllable break *after* an intervocalic consonant. In such cases we explicitly mark the syllable break with a period, e.g. item 31: *dam-ot* [dam.'ot] 'heavy'. In cognates from many other Philippine languages, there is a glottal stop in this position, so it appears that in Minangali the glottal stop has been deleted historically in these cases. Fourth, the phoneme transcribed in the Minangali orthography as <k> is normally realized as a glottal stop [ʔ] in Minangali. Cognates from many other Philippine languages have a /k/ in this position, so it appears that in Minangali the /k/ has evolved into /ʔ/. However, in item 68 it is a [k] that is produced: *sakkud* ['sakkud] 'horn'. It is not certain that this word is a borrowing. Fifth, the eth with lowering sign [ø̘] represents the interdental approximant, a rare speech sound found in about a dozen Philippine languages, including Kagayanen, Karaga Mandaya, Kalagan, Southern Catanduanes Bicolano, and several varieties of Kalinga (Olson and Mielke 2007).

2. WORDLIST.

<i>No.</i>	<i>English</i>		<i>Orthography</i>	<i>Phonetic</i>	<i>Notes</i>
1.	I		sakon	[ˈsaʔon]	
2.	you sing.; thou		sika	[ˈsiʔa]	
3.	he		siya	[siˈja]	
4.	we		ditaku	[ditaˈʔu]	inclusive
			dikani	[diʔaˈni]	exclusive
5.	you pl.		dikayu	[diʔaˈju]	
6.	they		dida	[diˈda]	
7.	this		anna	[ˈanna]	
8.	that		andi	[ˈandi]	
9.	here		sina	[siˈna]	
10.	there		sidi	[siˈdi]	
11.	who		ngai	[ŋai]	
12.	what		ngai	[ŋai]	
13.	where		dinu	[diˈnu]	
14.	when		kamaan	[ʔamaˈan]	
15.	how		inon	[iˈnon]	
16.	not		adipun	[adiˈpun]	negates verb or adjective
			dakampun	[daˈʔampun]	negates noun
17.	all		amin	[ˈamin]	
			losan	[loˈsan]	

18.	many	 adu	[a'du]	
19.	some	 udum	[u'dum]	
20.	few	 akit	[a'ʔit]	
		 poos	[ʔpoos]	
21.	other	 udum	[u'dum]	
22.	one	 ossaan	[os'saan]	
		 osa	[o'sa]	if counting
23.	two	 duwa	[du'wa]	
24.	three	 tu <u>l</u> u	[tu'ɔ̃u]	
25.	four	 opat	[o'pat]	
26.	five	 lima	[li'ma]	
27.	big	 dako <u>l</u>	[da'ʔoɔ̃]	
28.	long	 andu	[ʔandu]	
29.	wide	 a <u>l</u> wa	[ʔaɔ̃wa]	
30.	thick	 kopa <u>l</u>	[ʔo'paɔ̃]	
31.	heavy	 dagson	[dag'son]	
		 dam-ot	[dam.'ot]	
32.	small	 ban-og	[ban.'og]	
33.	short	 abobba	[abob'ba]	length
		 ansokba	[an'soʔba]	stature; morphology: an-sokba
34.	narrow	 ansu <u>l</u> pit	[an'suɔ̃pit]	morphology: an-su <u>l</u> pit

35.	thin	 ayyapit	[aj'japit]	thin; morphology: an-yapit
		 dagele	[dage'le]	skinny
36.	woman	 bubai	[buba'i]	
37.	man (adult male)	 la _l aki	[la'ða?i]	
38.	man (human being)	 tagu	['tagu]	
39.	child (a youth)	 abeng	[a'beŋ]	
40.	wife	 asawa'e bubai	[a'sawae buba'i]	literally 'fe- male spouse'; morphology: asawa-e bubai; mispronuncia- tion on second utterance
41.	husband	 asawa'e la _l aki	[a'sawae la'ða?i]	literally 'male spouse'; morphology: asawa-e la _l aki; hesitation on second utterance
42.	mother	 ina	[i'na]	
43.	father	 ama	[a'ma]	
44.	animal	 ayam	['ajam]	
45.	fish	 ugadiw	[uga'diw]	
46.	bird	 sissiwit	[sissi'wit]	
47.	dog	 asu	['asu]	
48.	louse	 kutu	['?utu]	
49.	snake	 u _l og	['uðog]	

50.	worm	🔊	batol _l	[ba'toŋ]	
51.	tree	🔊	kayu	['ʔaju]	
52.	forest	🔊	ginubat	[gi'nubat]	
53.	stick (of wood)	🔊	lasang	[la'saŋ]	
		🔊	tuŋu	[tu'ŋu]	piece of firewood
54.	fruit	🔊	buŋa	[bu'ŋa]	
55.	seed	🔊	bukol _l	[bu'ʔoŋ]	
56.	leaf	🔊	tubu	['tubu]	
57.	root	🔊	lamut	[la'mut]	
58.	bark (of tree)	🔊	uppak	['uppaʔ]	
59.	flower	🔊	tabbak	['tabbaʔ]	
60.	grass	🔊	lumun	[lu'mun]	
61.	rope	🔊	tali	[ta'li]	thick rope
		🔊	kiddong	['ʔiddoŋ]	generic rope
62.	skin (of a person)	🔊	kublīt	['ʔublīt]	
63.	meat (as in flesh)	🔊	laman	[la'man]	
64.	blood	🔊	da _l a	[da'ða]	
65.	bone	🔊	tung-a _l	['tuŋ.að]	
66.	fat (n.)	🔊	taba	[ta'ba]	
67.	egg	🔊	iplug	['iplug]	
68.	horn	🔊	sakkud	['sakkud]	possibly a loan
69.	tail	🔊	ipus	['ipus]	














70.	feather		dutdut	[dut'dut]	
71.	hair		buuk	[bu'uʔ]	
72.	head		u <u>l</u> u	[uʔu]	
73.	ear		inga	[iŋa]	
74.	eye		ata	[a'ta]	
75.	nose		ongol	[o'ŋoʔ]	
76.	mouth		sungad	[suŋad]	
77.	tooth (rather not molar)		baba	[ba'ba]	
78.	tongue		dela	[de'la]	
79.	finger nail		kukku <u>l</u> ung	[ʔuʔʔuʔuŋ]	
80.	foot		dapan	[da'pan]	bottom of foot
81.	leg		uki	[ʔu'ʔi]	includes foot
82.	knee		powog	[powog]	
83.	hand		ima	[ima]	includes arm
84.	wing		payak	[pa'jaʔ]	
85.	belly		buwang	[bu'waŋ]	
86.	guts		botak	[bo'taʔ]	
87.	neck		bagang	[ba'gaŋ]	
88.	back		odog	[o'dog]	
89.	breast		ba <u>l</u> ukung	[ba'ʔuʔuŋ]	
90.	heart		pusu	[pusu]	
91.	liver		agtoy	[agtoj]	

92.	drink	 uminum	[umi'num]	morphology: um-inum
93.	eat	 angan	[a'ŋan]	morphology: aN-kan
94.	bite	 kumtob	['ʔumtob]	morphology: um-kotob
95.	suck	 supsupan	[sup'supan]	morphology: supsup-an
96.	spit	 tumuppa	[tu'muppa]	morphology: um-tuppa
97.	vomit	 an-uta	[an.'uta]	morphology: an-uta
98.	blow (as wind)	 bumibbid	[bu'mibbid]	morphology: um-bibbid
99.	breathe	 umangos	[u'maŋos]	morphology: um-angos
100.	laugh	 mayokyok	[ma'joʔjoʔ]	morphology: ma-yokyok
101.	see	 ilan	[i'lan]	
102.	hear	 dongol	[do'ŋoɔ]	
103.	know (a fact)	 igammu	[igam'mu]	
104.	think	 somsomok	[somso'moʔ]	
105.	smell (sense odor)	 sungsungon	[suŋ'suŋon]	morphology: sungsung-on
106.	fear	 makimut	[maʔi'mut]	morphology: ma-kimut
107.	sleep	 nasuyop	[na'sujop]	morphology: na-suyop

















108.	live		matagu	[ma'tagu]	morphology: ma-tagu
109.	die		natoy	[na'toj]	morphology: na-toy
110.	kill		gongaton	[goŋa'ton]	morphology: gongat-on
111.	fight		ampatoy	[ampa'toj]	morphology: an-pa-toy
112.	hunt (transitive)		anganup	[aŋa'nup]	morphology: aN-anup
113.	hit		bikungon	[bi'u'ŋon]	morphology: bikung-on
114.	cut		sugaton	[suga'ton]	morphology: sugat-on
115.	split		bisakon	[bi'saʔon]	morphology: bisak-on
116.	stab (or stick)		duy-ukon	[duj.'uʔon]	morphology: duy-uk-on
117.	scratch (an itch)		gug-u	[gug.'u]	
118.	dig		amboka	[ambo'ʔa]	morphology: an-boka
119.	swim		ankiyat	[an'i'jat]	morphology: an-kiyat
120.	fly (v.)		tumayap	[tu'majap]	morphology: um-tayap
121.	walk		anadda _{lan}	[a _{lan} nadda'ðan]	morphology: aN-tadda _{lan}

122.	come		umali	[u'mali]	morphology: um-ali
123.	lie (as on one's side)		umbog	['umbog]	morphology: um-obog
124.	sit		tumukdu	[tu'mu?du]	morphology: um-tukdu
125.	stand		sumikad	[su'mi?ad]	morphology: um-sikad
126.	turn (change direction)		amballigus	[am,balli'gus]	morphology: an-balligus
127.	fall (as in drop)		naottag	[na'ottag]	morphology: na-ottag
128.	give		itton	['itton]	morphology: itod-on
129.	hold (in one's hand)		onnan	['onnan]	morphology: odon-an
130.	squeeze		posposon	[pospo'son]	morphology: pospos-on
131.	rub		apuwap	[a'puwap]	
132.	wash		kiwas	[?i'was]	
133.	wipe		punas	['punas]	
134.	pull		iyagwod	[i'jagwod]	morphology: i-agwod
135.	push		itukyud	[i'tu?jud]	morphology: i-tukyud
136.	throw		iballu	[ibal'lu]	morphology: i-ballu
137.	tie		itakod	[ita'?od]	morphology: i-takod
138.	sew		dait	['dait]	

139.	count		ambelang	[ambe'laŋ]	morphology: an-belang
140.	say		ibaga	[iba'ga]	morphology: i-baga
141.	sing		ankanta	[an'ʔanta]	morphology: an-kanta
142.	play		an-ay-ayam	[an.aj.'ajam]	morphology: an-ayayam
143.	float		anappak	[a'nappaʔ]	morphology: aN-tappak
144.	flow		an-ayus	[an.'ajus]	morphology: an-ayus
145.	freeze		bumatu	[buma'tu]	morphology: um-batu
146.	swell		lumtog	['lumtog]	morphology: um-lotog
147.	sun		init	['init]	
148.	moon		so _l ag	[so'ðag]	
149.	star		bituwon	[bi'tuwon]	
150.	water		danum	[da'num]	
151.	rain		udan	[u'dan]	
152.	river		tawwang	['tawwaŋ]	
153.	lake		baybay	[baj'baj]	
154.	sea (as in ocean)		baybay	[baj'baj]	
155.	salt		asin	[a'sin]	
156.	stone		batu	[ba'tu]	

157.	sand		gallis	[gal'lis]	
158.	dust		tapuk	['tapuʔ]	on trail
			gab <u>u</u>	[ga'bu]	in yard
159.	earth (as in soil)		luta	['luta]	
160.	cloud		libuu	[libu'u]	
161.	fog		nopnop	['nopnop]	
162.	sky		langit	['laŋit]	
163.	wind (as in breeze)		bibbid	[bib'bid]	
164.	snow		da <u>la</u> lu	[da'ɔ̃aɔ̃u]	
165.	ice		da <u>la</u> lu	[da'ɔ̃aɔ̃u]	
166.	smoke		asok	[a'soʔ]	
167.	fire		apuy	[a'puj]	
168.	ashes		dapo <u>l</u>	[da'poɔ̃]	
.	burn		sogob	[so'gob]	root form
170.	road		ka <u>l</u> sa	['ʔaɔ̃sa]	
171.	mountain		beleg	[be'leg]	
172.	red		dilaag	[di'laag]	
173.	green		lingsi	['lingsi]	
174.	yellow		ngela	[ŋe'la]	
175.	white		itak	['itaʔ]	
176.	black		ngitit	['ŋitit]	
177.	night		labi	[la'bi]	

178.	day (daytime)	🔊	al _g aw	[ʼaḏgaw]	
179.	year	🔊	tawon	[taʼwon]	
180.	warm (as in weather)	🔊	an-attub	[an.ʼattub]	morphology: an-attub
181.	cold (as in weather)	🔊	antungnin	[anʼtuŋnin]	morphology: an-tungnin
182.	full	🔊	napnu	[napʼnu]	
183.	new	🔊	kak-a _l a	[ʔaʔ.ʼaḏa]	
184.	old	🔊	dadaan	[dadaʼan]	
185.	good	🔊	amba _l u	[ambaʼḏu]	morphology: an-ba _l u
186.	bad	🔊	lawweng	[ʼlawweŋ]	
187.	rotten (as a log)	🔊	na _l pos	[ʼnaḏpos]	morphology: na-lopos
188.	dirty	🔊	kaissaw	[ʔaisʼsaw]	
189.	straight	🔊	nanggodong	[naŋgoʼdoŋ]	morphology: nan-godong
190.	round	🔊	natibuko _l	[na,ʔibuʼʔoḏ]	morphology: na(ti)-buko _l
191.	sharp (as a knife)	🔊	antadom	[antaʼdom]	morphology: an-tadom
192.	dull (as a knife)	🔊	bungdo _l	[buŋʼdoḏ]	
193.	smooth	🔊	an-imnas	[an.ʼimnas]	morphology: an-imnas
194.	wet	🔊	ambabasa	[ambabaʼsa]	morphology: an-CV-basa
195.	dry (adjective)	🔊	ammamaga	[ammamaʼga]	morphology: an-CV-maga
196.	correct (right)	🔊	ustu	[ʼustu]	

197.	near		adani	[ada'ni]	
198.	far		adayu	[ada'ju]	
199.	right (side)		dawinan	[dawi'nan]	
200.	left (side)		dawigi	[dawi'gi]	
201.	at		ad	[ad]	employed with location words
			atte	['atte]	
202.	in		ad	[ad]	employed with location words
			atte	['atte]	
203.	with (accompanying)				no entry
204.	and		kan	[?an]	
			kad	[?ad]	
			wok	[wo?]	
			ya	[ja]	
205.	if		nu	[nu]	
206.	because		te	[te]	
207.	name		ngadan	['ŋadan]	
169.	burn (intransitive)		nasgob	[nas'gob]	passive form ('to be burned'); morphology: na-sogob

REFERENCES

- BIRD, STEVEN, KAZUAKI MAEDA, XIAOYI MA, HAEJOONG LEE, BETH RANDALL, and SALIM ZAYAT. 2002. TableTrans, MultiTrans, InterTrans and TreeTrans: Diverse tools built on the Annotation Graph Toolkit. *Proceedings of the Third International Conference on Language Resources and Evaluation*. Paris: European Language Resources Association. <http://arxiv.org/abs/cs/0204006>.
- GORDON, RAYMOND G., ed. 2005. *Ethnologue: Languages of the world*, 15th edition. Dallas, TX: SIL International. <http://www.ethnologue.com>.
- HIMMELMANN, NIKOLAUS. 1998. Documentary and descriptive linguistics. *Linguistics* 36:161–195.
- IASA-TC03. 2005. The safeguarding of the audio heritage: Ethics, principles and preservation strategy, version 3. http://www.iasa-web.org/IASA_TC03/TC03_English.pdf.
- INTERNATIONAL PHONETIC ASSOCIATION. 1999. *Handbook of the International Phonetic Association*. Cambridge: Cambridge University Press.
- LADEFOGED, PETER. 2003. *Phonetic data analysis: An introduction to fieldwork and instrumental techniques*. Oxford: Blackwell.
- MACHLAN, GLENN, and KENNETH S. OLSON. 2008. Minangali (Kalinga) digital wordlist: Archival form. SIL-LCA-50319. SIL Language and Culture Archives, Dallas, TX.
- OLSON, KENNETH S., and JEFF MIELKE. 2007. Articulation of the Kagayanen interdental approximant: An ultrasound study. Paper presented at the Linguistic Society of America annual meeting, January 2007, in Anaheim, CA.
- PLICHTA, BARTEK, and MARK KORNBLUH. 2002. Digitizing speech recordings for archival purposes. http://www.historicalvoices.org/papers/audio_digitization.pdf.
- SIMONS, GARY F., KENNETH S. OLSON, and PAUL FRANK. 2007. Ngbugu digital wordlist: A test case for best practices in archiving and presenting language documentation. *Linguistic Discovery* 5(1):28–39. <http://journals.dartmouth.edu/cgi-bin/WebObjects/Journals.woa/2/xmlpage/1/article/314>.
- SWADESH, MORRIS. 1952. Lexico-statistic dating of prehistoric ethnic contacts: With special reference to North American Indians and Eskimos. *Proceedings of the American Philosophical Society* 96(4):452–463.
- SWADESH, MORRIS. 1955. Towards greater accuracy in lexiostatistic dating. *International Journal of American Linguistics* 21:121–137.

Kenneth S. Olson
ken_olson@sil.org

Glenn Machlan
glenn_machlan@sil.org