

Malaysian Journal of Social Sciences and Humanities (MJSSH)

Volume 3, Issue 2, April 2018

e-ISSN : 2504-8562

Journal home page:
www.msocialsciences.com

Root-Oriented Words Generation: An Easier Way Towards Dictionary Making for the Dusunic Family of Languages

Cosmas Julius Abah[1], Jane Wong Kong Ling[1], Anantha Raman Govindasamy[1]
[1]Centre for the Promotion of Language and Language Learning
Universiti Malaysia Sabah (UMS)

Correspondence: Cosmas Julius Abah (eddolabah47@gmail.com)

Abstract

Dictionary production is one of the most effective methods of preserving languages and cultures. The Dusunic Family of Languages (DFL) in Sabah, Malaysia would have welcomed the efforts to document their languages through dictionary production as there are still lacking of dictionary, vocabulary and phrase books. Furthermore, more than half of the languages in DFL are unwritten. However, making dictionary conventionally is tedious and time consuming. The Dusunic Family of Languages which are facing extinction threats do not have the luxury of time to wait for dictionary production via the conventional method. Hence, this study explores the use of a method called Root-Oriented Words Generation (ROWG) which is formulated based on spelling orthography of DFL to generate one and two-syllable words list. From the words list, root words registers were compiled which can then be used as database for dictionary production. Findings of this study showed that ROWG was able to generate an exhaustive word lists of DFL and compile a large volume of root words register in DFL. Hence, this study was able to highlight the feasibility and viability of using ROWG to produce root words register of DFL which could possibly reduce the time for dictionary production significantly. In future studies, it is recommended that the ROWG is extended to include more than two syllable words. This study showed the potentiality of ROWG to address the looming demise of DFL by providing a more efficient way of compiling root words for the purpose of making a dictionary.

Keywords: Dusunic family of languages, root words register, dictionary production, Dusunic roots-oriented words generation

Introduction

Sabah is known for its language diversity. This is indicated by the availability of 50 spoken languages of which, 33 are indigenous Austronesian family of languages. These 33 languages are from the Dusunic, Paitanic and Murutic family of languages, with Dusunic family of languages (DFL), as the largest one (Boutin, 2004). There are 17 languages in DFL, with eight being written languages (Bisaya Sabah, Tatana, Kimaragang, Lotud, Rungus, Tobilung, Kadazandusun and Kadazan Labuk Kinabatangan) and nine unwritten languages (Bisaya Brunei, Gana, Kadazan Klias River, Talantang, Kuijau, Papar, Minokok, Sugut Dusun and Dumpas).

Status of Dusunic Family of Languages

The 17 languages in DFL is facing an extinction threat. Based on the Expanded Graded Intergenerational Disruption Scale (EGIDS) that measures the status of a language based on development or endangerment, Table 1 shows that all 17 languages in DFL are facing various levels of endangerment. EGIDS measures the endangerment of these languages based on the scales of 3 (De factor National Working Language), 6B (threatened), 7 (shifting), 8a (moribund) (Lewis, Simons & Fenning, 2016).

Table 1: Dusunic Languages Classification and Status

No.	Language [ISO-693-3]	Written Level	EGIDS Scale	Endangerment Level
1	Bisaya Sabah [bsy]	Written	7	Shifting
2	Tatana [txx]	Written	6b	Threatened
3	Bisaya Brunei [bsb]	Unwritten	3	De facto National Working Language
4	Gana [gnq]	Unwritten	8a	Moribund
5	Kadazan, Klias River [kqt]	Unwritten	6b	Threatened
6	Kimaragang [kqr]	Written	6b	Threatened
7	Talantang [grm]	Unwritten	6b	Threatened
8	Kuijau [dkr]	Unwritten	6b	Threatened
9	Lotud [dtr]	Written	7	Shifting
10	Papar [dpp]	Unwritten	7	Shifting
11	Rungus [drg]	Written	6b	Threatened
12	Tobilung [tgb]	Written	6b	Threatened
13	Kadazan/Dusun [dtp]	Written	6b	Threatened
14	Minokok [mqq]	Unwritten	6b	Threatened
15	Sugut Dusun [kzs]	Unwritten	6b	Threatened
16	Kadazan, Labuk-Kinabatangan [dtb]	Written	7	Shifting
17	Dumpas [dmv] (unclassified)	Unwritten	6b	Threatened

A diligent search of published vocabulary, dictionary, wordlist and phrase books of the written languages in DFL has found a very limited volume of documentation as shown in Table 2. These words collection (WC) were the efforts from the early European compilers, Summer Institute of Linguistics (SIL), the Dusunic Cultural and Literary Organisations, Kadazandusun Language Foundation (KLF) and individual Dusunic speakers. Only 42 relevant WC were discovered.

Table 2: Words Collection of the Written Languages in DFL

Languages in DFL	Word Register	Phrasal Book	Vocabulary	Dictionary
Bisaya, Sabah [bsy]	1			
Tatana [txx]		2		
Kimaragang [kqr]		1		1
Lotud [dtr]				2
Rungus [drg]		1		5
Tobilung [tgb]		1		2
Kadazan Dusun [dtp]		7	3	11
Kadazan, Labuk-Kinabatangan [dtb]			1	

Contributing Factors to the Decline of DFL Usage

In general, modernization and globalization are often cited as the cause of development or endangerment of any language, which is also a factor in the decline use of DFL among native speakers. Other factors include the influence of the National Language Policy (NLP) which has exacerbated the endangerment of DFL with Malay being regarded as the language of education (Asmah, 1992). Malay is also the language of the media and the government (Lasimbang and Kinajil, 2004). Hence, the preference for Malay or more specifically, Malay Language Dialect has intruded the home domain abetted by parents who are concerned with better education and greater benefits of being assimilated into the Malaysian culture (Tumin, 2008; Wong, 2012). This was at the expense of using mother tongue. Mixed marriages are also another contributing factor to the decline in using mother tongue as code switching at home become prominent and the preference for the simpler and more convenient Sabah Malay dialect (Wong, 2000).

Development of the Kadazandusun Language

Kadazandusun has the most published word collection among the DFL. The Kadazandusun language (KD) was accepted as the official language for the Dusunic community and in 1995, KD was approved as the language taught under Pupil's Own Language (POL) programme in schools. KD which uses the Bunduliwan dialect has been taught in many schools in Sabah since 1997 (Smith, 2003; Lasimbang, 2000) and its domineering role has the effect of the minority languages being ignored.

According to the "Bunduliwan Agreement 1995", KD was supposed to be continuously enriched by other dialects that are found under the Dusunic and Paitanic group, which would imply that these enriching dialects would first be documented into wordlists or dictionaries in order to be an effective and meaningful enriching sources (Logijin, 2008). However, it seemed that it is unlikely to happen soon (Tangit, 2005). Thus, it had the effect of KD not being unanimously accepted as Dusun language takes precedent. For examples, speakers of Tangaa' do not find KD helpful to their mother-tongue maintenance and Rungus do not support KD as there are no Rungus words in KD (Tangit, 2005).
The focus on KD only as indicated by the support from the government and other language organizations such as SIL and KLF (Lasimbang & Kinajil, 2000), speakers of the Dusunic languages minority might continue to be ignored and eventually succumbing to extinction. To face the reality, speakers of these minority languages need to take responsibility of making their own dictionary. As Fishman (2001) said, "there can be no progress if speakers or the remaining speakers do not take on an active part, even to the point of relearning the language so that they can act as interlocutors….", these speakers need to be actively involved in documenting the language. However, being lexicographically handicapped, this would be a huge and difficult responsibility to undertake. The most daunting task in making a dictionary is in compiling the words collection of a language which may take several years. Hence, there is a need to explore methods or processes that could expedite the collection of words in order to make dictionary faster and less tedious. The language minorities have no luxury of time and are fast disappearing with its speakers growing old and dying with the passing time. Language revitalisation needs a more efficient way of collecting words so that documentation of the language can be expedited.

Root Words Register

The prerequisite of making a dictionary is in having the words in the source language, that is, the language in DFL and glossed with another known target language, normally English and/or Malay in the Malaysian context. Root-oriented entries dictionary are customarily done in the Austronesian family of languages' dictionary making (Newell, 1964). Efforts in the past such as Forschner (1992), Sokuroh (2013) and Abah (2015) advocated the use of root words register which is faster and less costly. This was agreed by Albright and Hatton (2007) who added that the efforts are less tedious with the direct involvement of the language community members.

Conventional Methods of Words Collection

Three conventional methods of collecting words are the Default Language Worker's Process (DLWP), the Text Corpus Method (TCM) and the Rapid Words Collection (RWC). DLWP is a layman's way of collecting words for the purpose of making a dictionary or vocabulary book. The method is slow, tedious and requires immersion with the native speakers who provide the words of the target language which are glossed with the source language (Antonissen, 1958; Samarin, 1967). TCM is a method of building a dictionary based on information that are attained from observation of text corpus of a language (Sinclair, 1991; Hurskainen, 2008). However, in the case of DFL which has lacking WC, TCM is not the best option for word collection. Another method called RWC was introduced by Moe (2003) as an attempt to revolutionise the task of words collection using a system of 1,800 semantic domains to capture words in a language workshop. These workshops were organised for the language community under the supervision of SIL (Moe, 2003; Manson, 2013). This method is effective but time consuming and costly as well.

Development of a New Dusunic Words Collection

The Root-Oriented Word Generation (ROWG) method is proposed as a one-time systematic generation of distinct DFL word list capturing DFL root words exhaustively in the words list. Philosophically, ROWG is an incidental discovery arising from the Rhyming Dictionary in Kadazan Tangaa'(RDKT) which was extended from the Kadazan Dusun Malay English Dictionary (Abah, 2011). The Rhyming Dictionary was intended as a reference by Kadazandusun song and poems writer to compose their creative writing. By prefixing consonants to the RDKT entries, other letter combinations were produced. ROWG used similar method to produce the word list. However, the word list contains both root words and nonsense syllables which need to be sifted by the well-versed speakers of the language to compile the root words register.

The conceptual foundation of ROWG is based on two main features of a language: orthography and syllable constituents. Orthography refers to the set of rules for writing a language which include spelling, hyphenation, capitalisation, word breaks, emphasis and punctuation. It is about letter identities and letter positions (Grainger & Hannagan, 2014). The orthographic structure of a written language identifies where particular letters appear within words, which sequence of letter is allowable and the pronounceability of words (Corcos & Willows, 1993). Orthographic coding is based on the formation of visual long-term memory representations of letters, letter patterns and sequence of letters which can be used to map spatially the temporal sequence of phonemes within words (Ehri, 2005). Hence the orthography identity of DFL can be determined and used to arrange letters based on syllabication procedures. The combination of orthography and syllabication rules for DFL in ROWG enables the generation of words of DFL systematically and exhaustively as a one-time requirement only. The word list can then be used to develop root words register either by comparing with available WC or verified by well-versed speakers of the language. In this study, the development of root words register is done by comparison with available WC and verification from the researcher for languages that he is familiar with.

Research Objectives

The main objectives of this study are:

i. To develop a Root-oriented Words Generation (ROWG) for the Dusunic Family of Languages;
ii. To generate word lists for Dusunic Family of Languages; and
iii. To demonstrate the compilation of root-words register for specific languages in the Dusunic Family of Languages using ROWG.

Research Methodology

This study undertook an inductive approach involving the search for patterns based on observations and developing the explanations of these patterns (Creswell, 2014). In this study, the patterns of words based on orthography and syllable constituents led to the generation of words list and compilation of root words register. This study also employed a qualitative method, using mainly secondary data sources in the form of existing WC to compare with the word list to develop the root words register. Primary data sources using the researcher's language knowledge was also used to enrich the root words register.

Initially, the determination of the number of syllables to be included in the study was done. An analysis of the WC indicated that from a total of 28,361 entries in 12 WC, one-syllable root words comprise of 2.66%, two-syllable words with 71.56%, three-syllable words with 21.01% and words with four or more syllables with 4.77%. Hence, this study focused on one and two-syllable root words, totalling 74.22% of the total root words, thus, indicating that most root words in DFL are captured by using ROWG. The generation of word list and root word register was aided by Microsoft Excel as a database to enable faster search and words comparison of root words.

Findings and Discussion

The findings of this study showed the DFL orthography, the syllable constituents for one-syllable and two-syllable words, the generated word lists and the collected root words register.

DFL Orthography

The combined DFL orthography for vowels phonemes, long vowel phonemes, diphthong phonemes, sequence of two vowels broken up with glottal stops, two vowels of different syllables, sequences of triphthongs, consonant phonemes and consonant combinations. Table 3 presents the combined DFL vowels phonemes. Five phonemes, /a/, /i/, /e/, /o/, and /u/ occur in all positions in a root word. However, Kimaragang, Rungus and Tobilung languages normally exhibit a basic four vowel phonemes, /a/, /i/, /o/, and /u/ but have also developed a fifth contrastive vowel phoneme, /e/ that occurs in all positions of a root word (Kroeger, 2008).

Table 3: Combined DFL Vowel Phonemes

Phoneme	Initial	Medial	Final
/a/	**a**buk [dtr] (dust)	b**a**gas [txx] (rice)	dar**a** [drg] (if only)
/e/	**e**tom [kgr] (black)	s**e**ra [drg] (when)	ad**e** [drg] (exclamation of surprise)
/i/	**i**rak [drg] (laugh)	t**i**mug [txx] (water)	dir**i** [dtb] (yonder)
/o/	**o**lim [drg] (deny)	k**o**dut [dtb] (pinch)	gay**o** [dtr] (big)
/u/	**u**sig [dtp) (bark)	c**u**rut [bsy] (cigar)	pas**u** [txx] (water jar)

DFL has five long vowels phonemes /aa/ /ee/ /ii/ /oo/ and /uu/ that occur in all positions both in open and closed syllable as shown in Table 4.

Table 4: Combined DFL Long Vowel Phonemes

Phoneme	Initial	Medial	Final
/aa/	**aa**k [dtp] [shriek of pig)	p**aa**t [bsy] (chisel)	ting**aa** [dtp] (look up)
/ee/	**ee**ti [tgb] (here)	p**ee**s [tgb] (knife);	t**ee** [tgb] (faeces)
/ii/	**ii**b (dtp] (pour out from kettle)	p**ii**d [dtp] (wipe)	tong**ii** [dtp] (mackerel)

/oo/	**oo**ng [dtp] (stammering)	t**oo**b [dtb] (high tide)	r**oo** [tgb] (chin)
/uu/	**uu**b [dtp] (to smoke)	b**uu**s [dtb] (gourd)	tov**uu** [dtb] (small basket)

DFL has six diphthongs phonemes /ai/ /au/ /iu/ /oi/ /ou/ and /ui/ that occur in all positions in open syllable only and split into different syllable if they end with a consonant (Robinson, 2005) as shown in Table 5.

Table 5: Combined DFL Diphthong Phonemes

Phoneme	Initial	Medial	Final
/ai/	**ai**t [dtp] (mention)	k**ai**no [dtb] (either); b**ai**no [dtp] (cloth)	bid**ai** [dtp] (hang to dry)
/au/	**au**so [dtb] (none)	g**au**k [bsy] (drunk)	an**au** [bsy] (tasteless)
/iu/	**iu**p [dtp] (sip)	b**iu**la [dtp] (violin); p**iu**k [dtb] (dimples)	kog**iu** [drg] (gorilla)
/oi/	**oi**gi [dtp] (pillar, post]	k**oi**so[txx] (having none); soira' [dtb] (when)	was**oi** [tgb] (adze)
/ou/	**ou**jan [dtb] (tired)	d**ou**so [dtp] (sin)	pop**ou** [drg] (ember)
/ui/	**ui**k [dtp] (shriek of pig)	s**ui**bo [dtp] (below); k**uil** [dtr](temple)	bal**ui** [txx] (bamboo post)

The occurrence of a sequence of two vowels broken up with a glottal stop is found in Lotud, Rungus and Kimaragang but may or may not occur in all positions in a root word as shown in Table 6. However, such occurrence was not observed in Bisaya Sabah, Tatana', Tobilung, Kadazandusun and Kadazan Labuk-Kinabatangan.

Table 6: Combined DFL Sequences of Two Vowels Broken Up with Glottal Stop

Sequence	Initial	Medial	Final
a'a	n**a'a** [drg] (to hand out)	d**a'a**y [dtr] (do not)	d**a'a** [drg] (if only)
a'i	??	w**a'i**g [dtr] (water)	??
a'o	??	??	??
a'u	??	b**a'u**r [dtr] (poem)	g**a'u** [drg] (paddle)
i'a	??	m**i'a**yas [dtr] (having sex);	??
i'i	**i'i**k [drg] (sound of mice)	b**i'i**k [drg] (bag)	??
i'o	??	l**i'o**w [dtr] (neck)	??
o'i	??	k**o'i**lo [dtr] (know)	??
o'o	**o'o**d [drg] (choke)	g**o'o**k [dtr] (gurgle)	r**o'o** [dtr] (chin)
o'u	??	k**o'u**ng [dtr] (pit)	??
u'a	**u'a**k [drg] (sound of vomiting)	l**u'a**k [dtr] (back of knee)	??
u'o	**u'o**g [drg] to consider)	p**u'o**d [dtr] (stump)	t**u'o** [drg] (mature)
u'u	**u'u**l [drg] (suckle)	l**u'u**ng [dtr] (flamingo)	??

The combined DFL has sequences of two vowels of different syllable of a·o, i·a, i·o, u·a and u·o, where each vowel retains its individual sound and occur in all positions as shown in the Table 7.

Table 7: Combined DFL Sequences of Two Vowels of Different Syllable

Phoneme	Initial	Medial	Final
a·o	**a·o**b [dtp] (eclipse)	p**a·o**k [dtp] (chaff)	ap**a·o** [dtp] cook vegetable)
i·a	**i·a**lo [dtb] (he/she)	d**i·a**n [bsy] (candle)	v**i·a** [dtb] (wild taro)
i·o	**i·o**n [txx] (live)	p**i·o**k [dtp] (silver)	p**i·o** [dtp] (how many)
u·a	**u·a**ni [dtp] (kind of mango)	t**u·a**t [tgb] (vien)	t**u·a** [tgb] (fruits)
u·o	**u·o**t [dtp] (ask, inquire)	g**u·ol** [kgr] (taro)	t**u·o** [bsy] (fish stupor)

The combined DFL triphthongs occur mostly at the final position in open syllable only as shown in Table 8. However, Lotud was observed as not adhering to this pattern in their root words formation.

Table 8: Combined DFL Triphthongs

Triphtongs	Initial	Final
aai	??	m**aai** [tgb] (do)
aau	**aau** [dtp] (to rob, snatch)	s**aau** [tgb] (underneath)
auu	??	p**auu** [dtp] (energetic, skilful)
iau	??	m**iau** [tgb] (alive)
iiu	??	h**iiu** [dtp] (to forget)
ioi	??	b**ioi** [tgb] (slender)
iou	**iou** [dtp] (older people)	k**iou** [dtp] (brinjal)
oii	??	p**oii** [dtp] (deligent)
ooi	??	g**ooi** [tgb] (fry without oil)
ooo	??	k**oo'o'** [tgb] (agree)
oou	??	h**oou** [dtp](voice)
oui	??	g**oui** [dtp] (fry without oil)
ouu	??	t**ouu** [dtp] (gourd)
uai	??	t**uai** [dtp] (rattan)
uau	??	s**uau** [drg] (pleasant)
uii	??	p**uii** [dtp] (rashes on the thigh)
uoi	??	d**uoi** [dtp] (mackerel)
uou	??	t**uou** [tgb] base of a container)
uui	??	k**uui** [txx] (cake)
uuu	??	k**uuu** [dtp] (defecate)

The combined DFL have 23 consonant phonemes /b/ /b̲/ /c/ /d/ /d̲/ /g/ /h/ /j/ /k/ /l/ /m/ /n/ /ng/ /ny/ /p/ /r/ /s/ /t/ /v/ /w/ /y/ /z/ and /'/. Phonemes /b/ /d/ /g/ /h/ /k/ /l/ /m/ /n/ /ng/ /p/ /r/ /s/ /t/ /w/ and /y/ occur in all positions. Phonemes /b̲/ /c/ /d̲/ /j/ ny/ /v/ and /z/ occur only in initial and medial positions, whereas phoneme /'/occurs initially, medially and in final positions. Table 9 shows these occurrences.

Table 9: Combined DFL Consonant Phonemes

Phoneme	Initial	Medial	Final
/b/	**b**abag [bsy] (tortoise)	gi**b**ang [dtb] (left)	koru**b** [bsy] (slice)
/b̲/	**b**akas [dtp] (wild pig)	a**b**al [dtp] (news)	-
/c/	**c**uka [dtr] (vinegar)	cu**c**ur [bsy] (fried banana)	-
/d/	**d**uvo [dtb] (two)	sa**d**ung [txx] (machete)	liyu**d** [kgr] (flood)

/d/	**d**abus [dtp] (outside)	a**d**at [dtp] (custom)	-
/g/	**g**iras [drg] (maggots)	ra**g**iw [dtr] (needle)	buro**g** [bsy] (grey cloud)
/h/	**h**uyan [dtp] (tired)	ku**h**at [dtp] (mushroom)	rowo**h** [bsy] (chin)
/j/	**j**aho [dtp] (casting net)	ka**j**ab [dtb] (shoulder)	-
/k/	**k**ilap [dtp] (flash)	ka**k**un [dtr] (towel)	piya**k** [dtr] (chick)
/l/	**l**uvang [dtb] (hole)	pa**l**ad [kqr] (palm of hand)	tigu**l** [dtb] (trumpet)
/m/	**m**aliw [dtr] (move)	lo**m**ok [dtb] (fat)	sodo**m** [dtr] (ant)
/n/	**n**iju [dtb] (coconut)	to**n**ok [txx] (calf)	suru**n** [txx] (wasp)
/ng/	**ng**onsis [kqr] (gums)	bo**ngi**t [kqr] (beard)	sayu**ng** [dtr] (crab)
/ny/	**ny**amuk [dtr] mosquito)	po**ny**u [dtr] (turtle)	-
/p/	**p**orit [kqr] (measles)	ka**p**ak [dtp] (axe)	tari**p** [kqr] (spleen)
/r/	**r**avo [drg] (padi stalks)	ko**r**ut [dtp] (spider)	tubi**r** [kqr] (small-pox)
/s/	**s**ulig [txx] (floor)	ru**s**od [tgb] (souls)	rongu**s** [tgb] (snout)
/t/	**t**igom [dtr] (oyster)	wo**t**is (calf of leg)	bala**t** [dtr] (sea cucumber)
/v/	**v**agas [dtp] (rice)	vi**v**ik [dtp] (whistle)	-
/w/	**w**otik [dtr] (colours)	da**w**at [dtp] (ink)	vazu**w** [drg] (to flatter)
/y/	**y**aki [dtp] (dandruff);	ka**y**ab [kqr] (shoulder)	para**y** [kqr] (rice)
/z/	**z**opos [dtp] (damp)	vu**z**ad [dtp] (unfasten)	-
/‘/	**i’**it [drg](to bite)	g**i’**ak [dtr] (shout)	iy**u’** [bsy] (shark)

The DFL syllable structure is CVC and CVVC (*Kementerian Pelajaran Malaysia,* 2008) and the syllable structure of (C)VCC or CCV(C) never occur (Robinson, 2005). Bating (2001) stated that consonant <ng> and <ny> are phonemes /ng/ and /ny/ respectively. Thus, there are no initial or final consonant clusters in written DFL because they do not allow more than one consonant in syllable onsets and codas. When consonant combinations occur medially, the first consonant forms the coda of the previous syllable, and the second consonant forms the onset of the latter. There are 16 permissible consonant combinations as shown in Table 10.

Table 10: Combined DFL Consonant Combinations

	Examples
bp	lu**bp**u [dtr] (to blow); lu**bp**u [kqr] (measles)
dj	lo**dj**ong [drg] (plant something among thickets)
ds	su**ds**or [kqr] (poke); pa**ds**ok [drg] (to step on)
dt	bo**dt**u [kqr] (heel); po**dt**ung [dtr] (big frog)
dz	ho**dz**od [drg] (overloaded); ra**dz**a [drg] (king)
gk	la**gk**au [kqr] (storehouse); si**gk**ak [dtb] (thin monkey)
mb	ga**mb**ar [kqr] (picture); ka**mb**ang [drg] (to swell)
mp	ta**mp**ak [dtr] (on top); to**mp**u [dtb] (pole cat)
nc	ca**nc**ang [bsy] (mince); ka**nc**u [bsy] (some kind of iron hook)
nd	ba**nd**ul [bsy] (sill); ta**nd**us [tgb] (spear)
ngg	la**ngg**uk [kqr] (uvula); ro**ngg**oy [dtr] (good)
ngk	ta**ngk**ap [tgb] (sheath for machete); ban**gk**ai [dtb] (corpse)
nj	ka**nj**ab [dtb] (grasshopper); ti**nj**ob [drg] (pierce, penetrate)
ns	ngo**ns**is [drg] (gum); pi**ns**an [txx] (cousin)

nt	a**nt**ad [dtb] (from); intad [txx] (from)
nz	jo**nz**o [dtp] (to trot); tanzag [drg] (to spread)

Syllable Constituents for DFL

The syllable constituents for the combined DFL indicated two types: the one-syllable constituents and the two-syllable constituents. Table 11 shows the combined DFL one-syllable constituents showing ten possible formations.

Table 11: Combined DFL One-Syllable Constituents

One-Syllable	Examples
V	i [dtp] (which); o' [dtp] (okay)
VC	ong [kqr] (if); oh [drg] (okay)
CV	ja [dtb] (us); nga' [dtp] (but)
CVC	gam [kqr] (blood); nung [dtp] (if)
VV	oo [kqr] (okay); ou [dtp] (I)
VVC	uub [dtp] (to smoke); aut [bsy] (belt)
CVV	pai [txx] (stingray); raa [dtb] (blood)
CVVC	ngaam [dtb] (correct); kaad [dtp] (palate)
VVV	aau [dtp] (snatch, rob)
CVVV	paii [bsy] (string gray); mooi [dtp] (so that)

Table 12 is the combined DFL two-syllable constituents showing that there are 33 possible formations.

Table 12: Summarised DFL Two-Syllables Constituents

Two-Syllable	Examples
V+V	i·a' [dtp] (to teach); i·a (you)
CV+V	si·a [bsy] (feel); du·o [txx] (two); si·a [dtp] (shirt); pu·a [dtb] (crab)
V+VC	u·ab [bsy] (yawn); u·ak [drg] (sound of vomiting); u·og [drg] (think); u·ot [kqr] (ask); a·ob [dtp] (eclipse)
CV+VC	ku·ar [bsy] (to mix up); si·am [txx] (ten); li·as [kqr] (cane); gi'ak [dtr] (shout); gu·ol [drg] (yam); pu·ok [dtp] (owl); tu·at [dtb] (vein)
V+VV	i·au [bsy] (alive); i·ou [dtp] (older people)
CV+VV	si·ou [dtp] (brave); su·au [drg] (pleasant); ki·ou [dtp] (brinjal); li·ou [dtb] (neck); bu·oi [txx] (length of time); lu·oi [dtb] (weak)
V+CV	o·ku [bsy] (I, me); a·ri [txx] (day); i·du [kqr] (to flee); a·wu [dtr] (ash); i·ra [drg] (advise); a·li [tgb] (healthy, well); u·gu [dtp] (shaky, loose); a·di [dtb] (younger brother)
CV+CV	wa·nu [bsy] (husband); ba·sa [txx] (read); gi·zu [drg] (tornado); pa·ku [tgb] (fern); so·ko [dtp] (bamboo shoots); pu·ri [dtb] (scabbies)
V+CVC	u·kun [bsy] (skirt); u·sing [txx] (cat); e·tom [kqr] (black); u·tow [dtr] (apron); u·yag [drg] (arouse); i·yong [tgb] (mother); u·tin [dtp] (can)
CV+CVC	ga·yuh [bsy] (big, huge); bi·lod [txx] (rice); pi·lat [kqr] (wound); da·yaw [dtr] (earlobe); ka·vuk [drg] (cloud); pa·rok [tgb] (husks of rice); ba·duk [dtp] (tinder)
VC+CV	ang·gu [bsy] (small wok); am·pu [txx] (to own); in·ta [drg] (eat raw); ong·go [tgb] (which one); an·ta' [dtp] (playful); in·si [dtb] (inch)
CVC+CV	ton·du' [bsy] (mark); jum·pa [txx] (meet up); bod·tu [kqr] (heel); gin·su [dtr] (lipstick); kan·sa [drg] (goose); kang·ku [tgb] (I say); kum·ba [dtp] (sago branches)
VC+CVC	an·tad [bsy] (from); in·tad [txx] (from); on·som [kqr] (sour); an·dad [dtr] (wait);

CVC+CVC	ang·gol [tgb] (can endure); on·tut [dtp] (fart); an·tad [dtb] (from) jang·gut [dtb] (beard); man·tad [dtp] (from); kong·kob [bsy] (coconut shell); gan·dum [txx] (corn); sang·kul [kqr] (hoe); pan·tow [dtr] (horse); pad·sak [drg] (to stamp on); pan·jar [tgb] (contract)
V+CVV	u·mou [bsy] (oil); a·kai [txx] (we); o·ngoi [kqr] (going); a·voi [drg] (weave); a·toi [tgb] (liver); a·hau [dtp] (to rob); u·voi [dtb] (sister-in-law)
CV+CVV	wa·sai [bsy] (waterfall); ka·dai [txx] (shops); wa·soi [kqr] (hatchet); ro·diu [dtr] (radio); ri·zau [drg] (can); pa·rai [tgb] (paddy); pu·nai [dtp] (pigeon); da·lai [dtb] (corn); ti·ngaa [dtp] (look up)
V+CVVC	a·maar [dtb] (expensive)
CV+CVVC	pu·gook [bsy] (toad); si·koon [dtp] (to coil); ti·koon [dtp] (a chair)
VC+CVV	um·pou [bsy] (purse); on·soi [txx] (good, nice); in·dai [dtp] (alone, isolated); am·pai [drg] (together); ang·gai [tgb] (intimate friend); ing·kaa [dtp] (like this)
CVC+CVV	som·pui [bsy] (puff slowly); kan·diu [txx] (eagle); lag·kau [kqr] (storehouse); bang·kai [drg] (corpse); lan·tui [dtp] (naked); lan·dui [dtb] (tadpole); tan·daa [dtp] (a cock); som·puu [dtp] (unite); tun·tuu [dtp] (finger); hing·kaa [dtp] (this way)
VC+CVVC	om·buud [dtp] (young coconut fruit)
CVC+CVVC	tom·puung [dtp] (coconut shell); tum·paak [dtp] (lightning and thunder); ting·gook [dtp] (snore); tim·baan [dtb] (jungle);
VV+CV	oo·di [dtp] (over there); uu·ti [dtb] (there (nearby); ai·so [dtp] (nothing)
VV+CVC	ou·rod [dtp] (round); ou·jan [dtb] (tired)
CVV+CV	tau·ri [bsy] (at the back or behind); koi·so [txx] (none); tou·gi [dtr] (bean sprout); toi·nu [dtp] (spy, observing); kai·no [dtb] (let's go); soo·mo [dtp] (meet up)
CVV+CVC	lou·son [txx] (hungry); sou·jor [dtr] (soldier); maa·han [dtp] (lazy); kaa·pad [dtb] (lying across); soo·min [dtp] (mirror); too·bong [dtp] (a well); mii·zan [dtp] (feel envious (of each other)
VVC+CV	oom·bo [dtp] (visit, walk); aan·sa [dtp] (fence)
CVVC+CV	naan·sa [dtp] (fenced)
VVC+CVC	aan·dak [dtp] (celebrated); aan·dang [dtp] (balance)
CVVC+CVC	guun·sing [dtp] (brass urn); daan·dang [dtp] (infringe (law)
VV+CVV	oi·dau [dtp] (very fast); aa·noi [dtb] (nearby but not in touching distance)
CVV+CVV	boi·bui [dtp] (temporary lamp); gou·gui [dtp] (stone in fruit)
CVV+CVVC	pii·saan [dtp] (oneness); koi·saan [dtp] (unity)

Generation of DFL Word List

The one- and two-syllable constituents in Table 11 and Table 12 were used to provide the spelling pattern of the respective one and two-syllable DFL root words based on the orthographic features presented from Table 3 to Table 10. The generation of the syllable constituents represent the complete word list, thus capturing all 74.22% of DFL root words. Table 13 shows the generation of DFL word list of one-syllable constituents of CVV. Other than this, word list based on one-syllable constituents of VVC and bVVC can also be generated.

Table 13: The Generated DFL Word List of One-Syllable Constituents of CVV

Nucleus VV	**CVV=(b-z) VV Generated Words**									
aa	baa	caa	daa	gaa	haa	jaa	kaa	laa	maa	naa
	ngaa	nyaa	paa	raa	saa	taa	vaa	waa	yaa	
	zaa									
ee	bee	cee	dee	gee	hee	jee	kee	lee	mee	nee
	ngee	nyee	pee	ree	see	tee	vee	wee	yee	zee

ii	bii	cii	dii	gii	hii	jii	kii	lii	mii	nii
	ngii	nyii	pii	rii	sii	tii	vii	wii	yii	zii
oo	boo	coo	doo	goo	hoo	joo	koo	loo	moo	noo
	ngoo	nyoo	poo	roo	soo	too	voo	woo	yoo	zoo
uu	buu	cuu	duu	guu	huu	juu	kuu	luu	muu	nuu
	nguu	nyuu	puu	ruu	suu	tuu	vuu	wuu	yuu	zuu
ai	bai	cai	dai	gai	hai	jai	kai	lai	mai	nai
	ngai	nyai	pai	rai	sai	tai	vai	wai	yai	zai
au	bau	cau	dau	gau	hau	jau	kau	lau	mau	nau
	ngau	nyau	pau	rau	sau	tau	vau	wau	yau	zau
iu	biu	ciu	diu	giu	hiu	jiu	kiu	liu	miu	niu
	ngiu	nyiu	piu	riu	siu	tiu	viu	wiu	yiu	ziu
oi	boi	coi	doi	goi	hoi	joi	koi	loi	moi	noi
	ngoi	nyoi	poi	roi	soi	toi	voi	woi	yoi	zoi
ou	bou	cou	dou	gou	hou	jou	kou	lou	mou	nou
	ngou	nyou	pou	rou	sou	tou	vou	wou	you	zou
ui	bui	cui	dui	gui	hui	jui	kui	lui	mui	nui
	ngui	nyui	pui	rui	sui	tui	vui	wui	yui	zui
ao	bao	cao	dao	gao	hao	jao	kao	lao	mao	nao
	ngao	nyao	pao	rao	sao	tao	vao	wao	yao	zao
ia	bia	cia	dia	gia	hia	jia	kia	lia	mia	nia
	ngia	nyia	pia	ria	sia	tia	via	wia	yia	
	zia									
io	bio	cio	dio	gio	hio	jio	kio	lio	mio	nio
	ngio	nyio	pio	rio	sio	tio	vio	wio	yio	zio
ua	bua	cua	dua	gua	hua	jua	kua	lua	mua	nua
	ngua	nyua	pua	rua	sua	tua	vua	wua	yua	zua
uo	buo	cuo	duo	guo	huo	juo	kuo	luo	muo	nuo
	nguo	nyuo	puo	ruo	suo	tuo	vuo	wuo	yuo	zuo

Table 14 shows the generation of DFL word list of two-syllable constituents of CV+hV. Other than this, word list based on two-syllable constituents of V+hVC, tV+hVC, CVn.tV, Vn.tVC, and tVn.tVC can also be generated.

Table 14: The Generated DFL Word List of Two-Syllable Constituents of CV+hV

Syllable V+hV	**CV+hV=(b-z) (V+hV) Generated Words**									
a·ha	ba·ha	ca·ha	da·ha	ga·ha	ha·ha	ja·ha	ka·ha	la·ha	ma·ha	na·ha
	nga·ha	nya·ha	pa·ha	ra·ha	sa·ha	ta·ha	va·ha	wa·ha	ya·ha	za·ha
e·ha	be·ha	ce·ha	de·ha	ge·ha	he·ha	je·ha	ke·ha	le·ha	me·ha	ne·ha
	nge·ha	nye·ha	pe·ha	re·ha	se·ha	te·ha	ve·ha	we·ha	ye·ha	ze·ha
i·ha	bi·ha	ci·ha	di·ha	gi·ha	hi·ha	ji·ha	ki·ha	li·ha	mi·ha	ni·ha
	ngi·ha	nyi·ha	pi·ha	ri·ha	si·ha	ti·ha	vi·ha	wi·ha	yi·ha	zi·ha
o·ha	bo·ha	co·ha	do·ha	go·ha	ho·ha	jo·ha	ko·ha	lo·ha	mo·ha	no·ha
	ngo·ha	nyo·ha	po·ha	ro·ha	so·ha	to·ha	vo·ha	wo·ha	yo·ha	zo·ha
u·ha	bu·ha	cu·ha	du·ha	gu·ha	hu·ha	ju·ha	ku·ha	lu·ha	mu·ha	nu·ha
	ngu·ha	nyu·ha	pu·ha	ru·ha	su·ha	tu·ha	vu·ha	wu·ha	yu·ha	zu·ha
a·he	ba·he	ca·he	da·he	ga·he	ha·he	ja·he	ka·he	la·he	ma·he	na·he
	nga·he	nya·he	pa·he	ra·he	sa·he	ta·he	va·he	wa·he	ya·he	za·he
e·he	be·he	ce·he	de·he	ge·he	he·he	je·he	ke·he	le·he	me·he	ne·he
	nge·he	nye·he	pe·he	re·he	se·he	te·he	ve·he	we·he	ye·he	ze·he
i·he	bi·he	ci·he	di·he	gi·he	hi·he	ji·he	ki·he	li·he	mi·he	ni·he
	ngi·he	nyi·he	pi·he	ri·he	si·he	ti·he	vi·he	wi·he	yi·he	zi·he

o·he	bo·he	co·he	do·he	go·he	ho·he	jo·he	ko·he	lo·he	mo·he	no·he
	ngo·he	nyo·he	po·he	ro·he	so·he	to·he	vo·he	wo·he	yo·he	zo·he
u·he	bu·he	cu·he	du·he	gu·he	hu·he	ju·he	ku·he	lu·he	mu·he	nu·he
	ngu·he	nyu·he	pu·he	ru·he	su·he	tu·he	vu·he	wu·he	yu·he	zu·he
a·hi	ba·hi	ca·hi	da·hi	ga·hi	ha·hi	ja·hi	ka·hi	la·hi	ma·hi	na·hi
	nga·hi	nya·hi	pa·hi	ra·hi	sa·hi	ta·hi	va·hi	wa·hi	ya·hi	za·hi
e·hi	be·hi	ce·hi	de·hi	ge·hi	he·hi	je·hi	ke·hi	le·hi	me·hi	ne·hi
	nge·hi	nye·hi	pe·hi	re·hi	se·hi	te·hi	ve·hi	we·hi	ye·hi	ze·hi
i·hi	bi·hi	ci·hi	di·hi	gi·hi	hi·hi	ji·hi	ki·hi	li·hi	mi·hi	ni·hi
	ngi·hi	nyi·hi	pi·hi	ri·hi	si·hi	ti·hi	vi·hi	wi·hi	yi·hi	zi·hi
o·hi	bo·hi	co·hi	do·hi	go·hi	ho·hi	jo·hi	ko·hi	lo·hi	mo·hi	no·hi
	ngo·hi	nyo·hi	po·hi	ro·hi	so·hi	to·hi	vo·hi	wo·hi	yo·hi	zo·hi
u·hi	bu·hi	cu·hi	du·hi	gu·hi	hu·hi	ju·hi	ku·hi	lu·hi	mu·hi	nu·hi
	ngu·hi	nyu·hi	pu·hi	ru·hi	su·hi	tu·hi	vu·hi	wu·hi	yu·hi	zu·hi
a·ho	ba·ho	ca·ho	da·ho	ga·ho	ha·ho	ja·ho	ka·ho	la·ho	ma·ho	na·ho
	nga·ho	nya·ho	pa·ho	ra·ho	sa·ho	ta·ho	va·ho	wa·ho	ya·ho	za·ho
e·ho	be·ho	ce·ho	de·ho	ge·ho	he·ho	je·ho	ke·ho	le·ho	me·ho	ne·ho
	nge·ho	nye·ho	pe·ho	re·ho	se·ho	te·ho	ve·ho	we·ho	ye·ho	ze·ho
i·ho	bi·ho	ci·ho	di·ho	gi·ho	hi·ho	ji·ho	ki·ho	li·ho	mi·ho	ni·ho
	ngi·ho	nyi·ho	pi·ho	ri·ho	si·ho	ti·ho	vi·ho	wi·ho	yi·ho	zi·ho
o·ho	bo·ho	co·ho	do·ho	go·ho	ho·ho	jo·ho	ko·ho	lo·ho	mo·ho	no·ho
	ngo·ho	nyo·ho	po·ho	ro·ho	so·ho	to·ho	vo·ho	wo·ho	yo·ho	zo·ho
u·ho	bu·ho	cu·ho	du·ho	gu·ho	hu·ho	ju·ho	ku·ho	lu·ho	mu·ho	nu·ho
	ngu·ho	nyu·ho	pu·ho	ru·ho	su·ho	tu·ho	vu·ho	wu·ho	yu·ho	zu·ho
a·hu	ba·hu	ca·hu	da·hu	ga·hu	ha·hu	ja·hu	ka·hu	la·hu	ma·hu	na·hu
	nga·hu	nya·hu	pa·hu	ra·hu	sa·hu	ta·hu	va·hu	wa·hu	ya·hu	za·hu
e·hu	be·hu	ce·hu	de·hu	ge·hu	he·hu	je·hu	ke·hu	le·hu	me·hu	ne·hu
	nge·hu	nye·hu	pe·hu	re·hu	se·hu	te·hu	ve·hu	we·hu	ye·hu	ze·hu
i·hu	bi·hu	ci·hu	di·hu	gi·hu	hi·hu	ji·hu	ki·hu	li·hu	mi·hu	ni·hu
	ngi·hu	nyi·hu	pi·hu	ri·hu	si·hu	ti·hu	vi·hu	wi·hu	yi·hu	zi·hu
o·hu	bo·hu	co·hu	do·hu	go·hu	ho·hu	jo·hu	ko·hu	lo·hu	mo·hu	no·hu
	ngo·hu	nyo·hu	po·hu	ro·hu	so·hu	to·hu	vo·hu	wo·hu	yo·hu	zo·hu
u·hu	bu·hu	cu·hu	du·hu	gu·hu	hu·hu	ju·hu	ku·hu	lu·hu	mu·hu	nu·hu
	ngu·hu	nyu·hu	pu·hu	ru·hu	su·hu	tu·hu	vu·hu	wu·hu	yu·hu	zu·hu

Compilation of the Root Words Register

From the DFL word lists, the compilation of root words register for the languages in DFL was made possible as shown from Table 15 to Table 20 for Rungus, Kadazan Penampang, Kadazan Papar, Tobilung, Lotud and Kadazan Labuk-Kinabatangan.

Table 15: The Rungus Root Words Register

Root Words	Meanings
aha	misfortune
ahab	to flirt
ai	exclamation of excitement
anta	ready to strike
antad	from
antak	to fill up a container
bu'ut	paddy about to flower
dahu	to go downstair
ihad	to cry
ihib	to pour out
ihud	footmarks; footprint

i'ik	sound of mice
ontok	right place
ontong	to marry
ontut	fart
o'od	choke
o'ok	sound of some one being strangled
paha	honey from bees
raha	blood
tihi	drip; leak
tihis	to filter
uhad	to press down
uhol	to squirm weakly
untu	top part of trees
u'ul	to suckle
vaha	boil
zuhu	willingness to share

Table 4.16: The Kadazan Penampang Root Words Register

Root Word	Meanings
a'a	no
aak	shriek of pig
aha'	defeat
ahom	deep
ahung	entice
ais	ice
antob	obstruct
baas	bas
bai	compete
bantu	threshold
biak	(whitist) frog
bii'	see, look
biid	small snail
gua’	(to) shake, move
hia	that
hinta'	appear, come out
hontu	smell of urine
ia'	teach
iad	cry
ohig	put barrier
ohim	deny
poo	thigh
tantang	put fencing to fasten a mad man
tantop	frame
tia	disobedient
tiha'	lick
tontong	fascinated, amaze
uhus	to slide down
unta'	camel
uud	becoming loose
uung	fall (of fruit)
uus	chew something until it become tasteless

Table 17: The Kadazan Papar Root Words Register

Root Words	Meanings
ahab	ablaze, aflame
ahad	wing
binti	kick leg to make one fall
bius	anesthesia
biut	twisted
boud	a bulge
buil	cheek pouch
gia	please
gua	(to) move
guhi	coming back
hinta	to appear
hontu	pungent
iad	similar
iho	identify
ihob	vomit
ihong	caterpillar
kohu	sad looking
ngua	moo
nuo	take
saha	mistake
suun	to carry on head
tohu	three
tontom	crossbeam
uap	vapour
uhu	head
uhud	put in order
uhun	people
uik	shriek of pig
uin	to steer a canoe
vahu	eight
vuhu	feather, hair
yaha'	defeat
yuo	pain, hurt

Table 18: The Tobilung Root Words Register

Root Words	Meanings
aak	to guide one to walk
aas	to check and take the product
antak	raining
antas	short cut
baa	yes, okay
baas	bas
beet	comment, admonition
biik	young rat
binti	daughter of (followed by father's name)
boom	bomb
boop	bulb
buul	ball
buun	small container for chewing paraphernalia

buut	boat
doo	his or hers
gia	empathic particle
iis	expression of annoyance
intob	calculate
koi	we, but not you
o'o	denoting agreement with speaker
tantan	to forge, to smith;
tantu	certain, sure
tee	defecating
tuntung	percussion stick
tuu	change clothes
woo	beware! move!

Table 19: The Lotud Root Words Register

Root Words	Meanings
dahu'	recede (of water)
gi'i	tear
ihad	cry
ihi	bray
ihob	cough, sneeze
ihum	search
ihup	sip
intang	look at
kahu	scratch
lanta'	eyes open
ohun	charcoal
ontim	sea urchin
ontut	fart
paha	rack
paha'	juice
pahu'	mango
ruhu'	semen
tahum	birth mark
ta'i	feaces
ta'u	maybe
tihis	dribble
tontob	enough
tontong	17th day of lunar calendar
tuhap	snatch
tuntug	pour
tuntul	shell
tu'u	dry
uhab	shock
uhat	nerve
uhok	words or sounds uttered as a magic spell
uhot	ask

Table 20: The Labuk-Kinabatangan Kadazan Root Words Register

Root Words	Meanings
aa	that (far from the speaker but near to the listener
antad	from
anting	(anting-anting) earring
au	no, not
buuk	book
buur	drill
buus	gourd
daa	indicates certainty
dii	hey
kou	you (plural)
moi	somewhat, seems
naa	what, ah, oh
ontok	as long as, when
oo	okay, yes
paa	broth, gravy
poo	thigh
raa	blood
suu	command
tantu	decide
tii	tea
tonton	female genital
tuntu	extreme end
tuu	dry
uun	motor horn

Implications

The findings showed that there is an advantage of having DFL word lists which can be used as a reference to sift through the one- and two-syllable constituents and find root words in a specific language in DFL. The use of ROWG provides a more convenient way of comparing the word list with any word collection in DFL. It can also be used in RWC where well-versed speakers can go through the word list to sift root words in their respective language from nonsense syllables. Hence, it speeds up the process of compiling root words in any language of DFL. By extension, speakers of the other unwritten languages of Minokok, Dusun Sugut, Gana, Kuijau, Papar and Bisaya Brunei would also find the DFL word lists applicable to their respective language as it is mentioned in SIL Ethnologue (2016) that these unwritten languages have intelligibility and lexical similarity, in varying degree, with the written languages, especially with their neighbouring written languages.

The application of ROWG does not mean that other conventional methods can be replaced. On the contrary, ROWG should be used with other methods such as TCM and RWC for more effective yield of root words. The success in compiling root words registers in DFL demonstrates that ROWG is also applicable in other Austronesian family of languages.

Recommendations

The ROWG development has given opportunity to future researchers some fruitful research works for which this research has laid the foundation to researching in the following areas:

i. To extend the DFL word list to include three syllables or more;
ii. To explore the phonology of the words in the root word register to enrich the glossary of words in terms of spelling and phonetics;

iii. To develop some kind of filtering system to screen the word lists of nonsense morphemes;
iv. To design software for the ROWG generation of word lists; and
v. To make learning DFL root words interesting, innovative and fun by designing some kind of words game based on root words.

Conclusion

The development of ROWG is considered timely as it gives hopes to passionate speakers of DFL to compile root words register on their own despite being lexicographically handicapped. It is hoped that this would motivate more DFL speakers to produce dictionary or other published materials to preserve and perpetuate their mother tongue without being too dependent on language organisations. Hence, this could arrest or at least slow down the demise of the languages in DFL. This also provides an alternative for the unwritten languages in DFL to at least salvage some of the language before the native speakers are gone from the community. Assured that these languages are preserved to posterity would be a source of great joy and pride of the community. In addition, this would enhance the prestige of the language as well as improve the well-being and quality of life of the community at the same time.

References

Abah, C. J. (2011). *Rhyming Dictionary in Kadazan Tangaa'*. Kota Kinabalu, Sabah, Malaysia: Rhymas Media Production.

Abah, C. J. (2015). *Kadazan Dictionary (in Tangaa' Papar) Kota Kinabalu*, Sabah, Malaysia: Rhymas Media Production.

Albright, E. & Hatton, J. (2007). Chapter 10. WeSay, a tool for engaging communities in dictionary building. In D. Victoria Rau and Margaret Florey (eds). 2007. *Documenting and Revitalizing Austronesian Languages*, 189-201. Honolulu: University of Hawaii Press.

Antonissen, A. (1958). *Kadazan Dictionary and Grammar: Kadazan-English Dictionary section*. Canberra, Australia: Government Printing Office.

Asmah, H. O. (1992). *The Linguistic Scenery in Malaysia*. Kuala Lumpur. Dewan Bahasa dan Pustaka

Boutin, M. (2004). Languages of Sabah. In Asmah Haji Omar, *The Encyclopedia of Malaysia: Language and Literature*, 8, 38-39. Kuala Lumpur: Archipego Press.

Corcos, E., & Willows, D. (1993). The role of visual processing in good and poor readers' utilisation of orthographic information in letter strings. In S. Wright & R. Groner (Eds.), *Facets of dyslexia and its remediation*. Amsterdam: Elsevier

Creswell, J. D. (2014). *Research design: qualitative, quantitative, and mixed methods approaches*. Thousand Oaks, California: SAGE publications.

Ehri, L. (2005). Learning to read words: Theory, findings, and issues. *Scientific Studies of Reading, 9*(2), 167–188.

Fishman, J. (2001). *Can Threatened Language be saved, Reversing Language Shift, Revisited: 21st Century Perspective*. Clevedon: Multilingual Matters.

Forschner, T. A. (1992). *Register of roots in the Rungus dialect: Pongumpugan dot ongopuun dit pongoretan do Rungus*. Kudat

Grainger, J. & Hannagan, T. (2014). What is special about orthographic processing? *The Architecture of Writing Systems, 17*(2), 225-252.

Hurskainen, A. (2008). *SALAMA Dictionary Compiler - A Method for Corpus-Based Dictionary Compilation*, Institute for Asian and African Studies, Box 59 FIN-00014 University of Helsinki, Finland arvi.hurskainen@helsinki.fi; Technical Reports in Language Technology Report No 2. Online: http://www.njas.helsinki.fi/salama/salama-dictionary-compiler.pdf accessed on 20/4/2016

Kementerian Pelajaran Malaysia. (2008). *Puralan Boros Kadazandusun Id Sikul*. Putrajaya: Perpustakaan Negara Malaysia.

Kroeger, P. (2008). *Vowel Harmony in Kimaragang as a Lexical Rule*.

Kroeger, P.R. (1989).*Trilingual Phrase Book: Dusun Kimaragang-Bahasa Malaysia-English*. Kota Kinabalu: United Dusun Sabah Association

Lasimbang, R. (2004). To Promote the Kadazandusun Languages of Sabah. *ABD, 34*(2), Kota Kinabalu, Sabah: Kadazandusun Language Foundation(KLF)

Lasimbang, R. and Kinajil, T. (2000). Changing the Language Ecology of Kadazandusun: The Role of the Kadazandusun Language Foundation. *Current Issues in Language Planning*

Lewis, M. P, Simons, G. F. & Fennig, C.D. (eds.). (2016). *Ethnologue: Languages of the world* (19th edition). Dallas, TX: SIL International.

Logijin, S. (2008). A Case Study on the Use of Kadazandusun Language of Malaysia. A paper presented at the *Regional Consultative Workshop on "Using the Mother Tongue as Bridge Language of Instruction in Southeast Asian Countries: Policy, Strategies and Advocacy"*. Bangkok, Thailand,19-21 February 2008.

Manson, K. (2013). *Rapid Words Collection and Historical Linguistics*. Retrieved from https://www.google.co.nz/?gws_rd=ssl#q=Ken+Manson%2C+2013:+Rapid+Words+Collection+and+Historical+Linguistics, accessed on 20/4/2016.

Moe, R. (2003). Compiling Dictionaries Using Semantic Domains. *Lexikos 13* (AFRILEX-reeks/series 13:2003), 215-223. Grahamstown, South Africa

Newell, L.E. (1964). *Independent Clause Types of Batad Ifugao*. Oceanic Linguistic, *3* 3: 171-99).

Robinson, L. C. 2005. *A Sketch Grammar of Tindal Dusun*. Department of Linguistics University of Hawai'i at Mânoa Honolulu 96822

Samarin, W. T. (1967). *Field linguistics*. New York: Holt, Rinehart & Winston.

SIL Ethnologue. 2016. *Ethnologue: Language of the World*, retrieved from: https://www.ethnologue.com/

Sinclair, J. McH (1991). *Corpus, Concordance, Collocation*. Oxford University Press

Smith, K.J. (2003). Minority Language Education in Malaysia: Four Ethnic Communities Experiences. *International Journal of Bilingual Education and Bilingualism*. 6(1), 52-65.

Sokuroh, K. J. (2013). *Momogun-Rungus Rootwords*. Kota Kinabalu, Sabah: Persatuan Penulis Rungus Sabah (PENERUS).

Tangit, T. M. (2005). *Planning Kadazandusun Sabah, Malaysia: Labels, Identity and Language*. University of Hawai'i

Tumin, L. (2008). *Language attitudes of Kadazandusun people toward their mother tongue*, Project Paper, Universiti Malaysia Sarawak

Wong, J. K. L. (2000). *The Sabah Malay Dialect: Phonological Structures and Social Functions*, Kota Kinabalu: Penerbit UMS.

Wong, J. K. L. (2012). The Sabah Malay Dialect Phonological Differentiation in Social Context, *PhD Thesis*, Universiti Malaya, Kuala Lumpur.